THE UNIVERSITY OF ALBERTA

THE DEVELOPMENT AND EVALUATION OF A PROCESS APPROACH TO THE TEACHING OF JUNIOR HIGH SCHOOL SCIENCE

by

© ERIC MOKOSCH

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN

PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

FALL, 1969

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "The Development and Evaluation of a Process-Approach to the Teaching of Junior High School Science" submitted by Eric Mokosch in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

In the past few years a strong movement has become noticeable to change the emphasis in the teaching of science. Recent emphases have been on the teaching of science according to the definition of science - recognizing both its substantive and syntactical nature.

The study reported here is an attempt to incorporate the syntactical nature of science into science teaching, while at the same time conceding the importance of the substantive or content dimension.

A comparison was made among four treatment groups of their performance on an achievement test and on a number of measures specifically designed to measure "process" objectives. The treatments varied according to the emphasis they placed on the process dimension <u>and</u> the extent of involvement of teachers in the production of the experimental curriculum materials. Treatments A and B received materials based on a theoretical framework incorporating both the content and process dimensions of science, the latter being defined by "An Inventory of Processes in Scientific Inquiry". Processes of scientific inquiry were conscientiously integrated into the teaching materials for both groups. The teachers of treatment group B, however, did in no way participate in the construction of this experimental curriculum while those in Treatment A did. In Treatment C, the same subject matter was taught in a more or less traditional fashion while Treatment D served as a control group for testing.

Due to a lack of existing instruments measuring the process

dimension, the TOUS being the only suitable one found, three process instruments were constructed and validated especially for this study.

The major objectives of the study were:

1. The development of a curriculum recognizing both the conceptual and process nature of science.

2. The measurement of the growth in understanding of the processes of scientific inquiry along with measurement in gains on a more traditional measure of achievement.

3. The overall evaluation of the feasibility of a process approach to the teaching of junior high school science.

The following conclusions were drawn from the investigation:

1. A workable curriculum, recognizing the dual nature of science and at the same time emphasizing the process dimension, can be constructed.

2. All treatment groups, except the control group, performed equally well on the conventional measure of achievement.

3. The process measures as a whole failed to discriminate among treatments in the anticipated direction.

4. The expected high correlation between the TOUS and the process measures did not materialize.

The investigation has highlighted the need for longitudinal studies based on similar designs and, perhaps, the need for a continuing search for more sensitive process measures which measure process-oriented objective with greater statistical confidence.

ACKNOWLEDGEMENTS

The writer wishes to express his appreciation to the many persons who assisted in the completion of this study. In particular, I wish to express my appreciation to the members of my committee:

Dr. M. A. Nay, my supervisor, whose inspiration and continuing guidance have been vital factors in the inception and completion of this study;

Dr. S. B. Woods, whose continued interest and questioning helped clarify ideas;

Dr. D. W. R. Wilson, who made valuable suggestions for revision and lent moral support;

Mr. I. Hastings, who showed interest by taking time off from his busy schedule for a number of consultations;

Dr. R. S. MacArthur and Dr. T. Maguire, who helped conceptualize the statistical design and analyses.

Special acknowledgement is also made to the various teachers who participated in this project without reservation; the teachers and students from Hardisty Junior High, Vernon Barford, Avalon, Westminster, Strathearn, and Wellington.

A special acknowledgement is also expressed to other members of the Edmonton Junior High School Science Project, especially to Dave Thomas, Dave Powley, Henry Visscher, Don Whitlock and Rudy Melnychuk.

The writer also gratefully acknowledges the encouragement, patience and sacrifice given by his wife.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

I. THE BACKGROUND OF THE STUDY

In the mid-1950's, the public became very vocal about its dissatisfaction with existing science curricula and many prominent personalities in science and education were sufficiently concerned to examine the situation in some detail. This concern was reflected in the establishment of a number of national science curricula in the United States such as PSSC (Physical Science Study Committee) (47), CHEM Study (Chemical Education Material Study) (12), CBAC (Chemical Bond Approach Committee) (11), ESCP (Earth Science Curriculum Project (18), and BSCS (Biological Science Curriculum Study) (5) to mention only a few.

These science curricula had the support and guidance of many eminent scientists and leading science educators and they made many claims of alleviating the shortcomings of traditional science education. Influenced by theorists like Schwab (54), Brandwein (8) and Bruner (9), these studies advocated the selection of only the pervasive and encompassing concepts and ideas, the structuring of these for easier and more permanent learning, and greater emphasis in the teaching and learning of science on the way science knowledge is discovered and

and evolved.*

Many well-intended objectives were stated for these projects. An illustration of the philosophy underlying most of the new courses was found in the foreword of the CHEM Study text:

> A clear and valid picture of the steps by which scientists proceed is carefully presented and repeatedly used. . . Heavy reliance is placed upon laboratory work so that chemical principles can be drawn directly from student experiences. Not only does this give a correct and nonauthoritarian view of the origin of chemical principles, but it gives maximum opportunity for discovery, the most exciting part of scientific activity. . . The elements of scientific activity are immediately displayed, including the role of uncertainty. . . chemistry is gradually and logically unfolded, not presented as a collection of facts, dicta and dogma (13).

Unfortunately, however, these programs did little more than bring traditional content up to date. As far as the processes of science are concerned, if these were mentioned in the objectives at all, they appeared to be given only lip service in the teaching; aspects such as discovery learning in the laboratory were "tacked on" to ordinary content coverage at best. Understandably, then, even less was done in the way of evaluating for pupils' understanding of the processes of science.

How can a curriculum be built recognizing both the conceptual structure and the processes of inquiry in a given subject area? Parker and Rubin see the following tasks confronting the curriculum

---

*This dimension of science is usually referred to as the "process or the "processes" of science or scientific inquiry.

worker:

1. A retooling of subject matter to illuminate base structure, and to insure that knowledge which generates knowledge takes priority over knowledge which does not;

2. An examination of the working methods of the intellectual practitioner: the biologist, the historian, the political scientist, for the significant processes of their craft, and the use of these processes in our classroom instruction: . . . (44).

Even if the priority and primary importance of the first task is recognized, the translation of the second task into specific operational detail, thereby making it useful for the classroom teacher, is still a difficult one. In an effort to overcome this obstacle, the Edmonton Junior High School Science Project* was set up in Edmonton to investigate this problem, concentrating in particular on the second task.

The immediate product of the research and deliberations was <u>An Inventory of Processes in Scientific Inquiry</u>** which was a descriptive representation of what scientists do when they pursue science. At the same time, it was also a very useful instrument for structuring the teaching of scientific topics. Thus this versatile instrument not only allowed for the conceptual structuring of each topic taught, but also

---

*The project is directed by M. A. Nay, Department of Secondary Education, University of Alberta, and co-directed by R. Melnychuk, Junior High School Science Supervisor, Edmonton Public School Board. It is a cooperative research and development endeavour involving junior high school science teachers and university personnel (including the author).

**This particular Inventory will hereafter be simply referred to as "Inventory". See Appendix A for a copy. An earlier draft of the "Inventory" was included in an article written by Nay (41) in the ATA Magazine.

directed the teacher to consciously plan and teach for the processes in scientific inquiry.

Assuming for the moment that curricula can be constructed which reflect both the product and processes of inquiry of science, how is student attainment or growth along these dimensions to be evaluated? An examination of science tests in the last fifty years dealing with the cognitive domain showed a clear emphasis on the testing for the achievement of content-objectives; the determination of the attainment of objectives related to the process dimension being virtually neglected. This point of view is supported by Lizonbee (37) when he stated that most of the testing done in schools today is an evaluation of a student's ability to remember descriptive information, and by Watson's analysis of research on teaching science:

> The almost universal emphasis upon gains in scores upon achievement tests of limited scope is most alarming. This emphasis implies that the primary function of a teacher is to develop in the pupil only the small range of academic behaviors, and usually rather minor ones, which are required by the tests used (64).

What appears to be required is a more comprehensive system of evaluation which also attempts to measure growth along the process dimension, at the same time being as consistent with the way students are instructed as possible. Prior to the development of their TOUS*

---

*This test was designed to measure a student's understandings about the processes of scientific inquiry, the scientific enterprise, and the characteristics of scientists.

(Test On Understanding Science), Cooley and Klopfer decried the paucity of evaluation instruments and techniques in this area:

> However, while a large variety of tests have been prepared to measure student achievement in the facts and principles of science, no adequate instrument has yet been constructed to assess the extent to which the important instructional outcomes of understanding science and scientists has been achieved (16).

Once again, the versatility of the Inventory produced by the Edmonton Junior High School Science Project is apparent. The fact that the identified processes are stated in specific behavioural or operational terms allowed the construction of "process-measures"* in the familiar paper-and-pencil mode. However, in order to be as consistent as possible with the student's experiences during learning, and in order to avoid a confounding reading factor, it was felt that the stimuli presented to the student in the test situation should be as close to those encountered in the learning situation as possible. Schalock, Beaird and Simmons expressed this point of view when they stated:

> . . .it follows that one must look beyond traditional paper and pencil tests, with the reliance on words and related symbols for the presentation of the stimulus reactions to be used in tests predicting complex behavioural phenomenon. It follows also that as test stimuli are made more life-like, not only will the responses be more representative of life responses, but also the potentiality exists for variables affecting complex behaviour to become available and thereby, enhance predictions. Thus tests which utilize stimuli corresponding closely to

---

*See pages 8 and 9 for a description of these measures.

> actual situations should prove to be more effective predictors, because simultaneously they will (1) call forth responses that are more like those occurring in life situations and (2) make available for measurement a greater number of factors affecting complex behaviour (52).

Accordingly, a film loop format was chosen in order to present stimuli to the student in the test situation. Film loops allowed the presentation of stimuli which were very similar to those encountered during the learning phase of instruction.

Basing the testing on the Inventory not only allowed testing which was specific to the content investigated, but also gave the opportunity to measure gain in understanding the processes in scientific inquiry.

Three such "process-measures" were developed in an effort to incorporate Bruner's (9) ideas on transfer, both specific and general. Visualized were successive levels of knowing and understanding the processes of scientific inquiry. The three tests attempted to measure how well a student understood the processes in an investigation he performed, how well he was able to apply them in a parallel problem situation not encountered before, and finally, how well he was able to use them in a general science content area not directly related to that studied.

II. THE PROBLEM TO BE INVESTIGATED

This study, as the title suggests, was primarily concerned with finding possible answers to the following question: Can process in science be taught by a specially constructed curriculum; and if so,

can growth along the indicated dimensions be effectively measured? This question was studied in the context of comparison of four treatment groups. The groups were defined according to the type of instruction they received. Comparison among the groups aimed at discovering if differences accrued from the type of instruction received. The effect of the various treatments was investigated in order to determine to what extent the pupils' posttest scores on a variety of measures was determined by the instruction they encountered.

Definitions

Generally, each of the following terms is explained where it is first used in the study, but, for clarity, the following definitions are presented at this point:

Inventory refers to An Inventory of Processes in Scientific Inquiry developed by the Edmonton Junior High School Science Project under the direction of M. A. Nay.

Achievement test score refers to the number of items a student had correct on a specially constructed test to measure achievement in the content area of "Matter and Energy"*.

Cooperative science test score (COOP test score) refers to the number of items a student had correct on the Cooperative Science Test-

*See Appendix B for a copy of this test.

General Science - Form B.

Experimental Curriculum refers to the curriculum developed by the Edmonton Junior High School Science Project on the topic of "Matter and Energy". A special effort was made to integrate as much as possible, the processes of scientific inquiry with the content. The "Investigation" was the basic teaching and learning unit in this course.

Process measure score (Part I) refers to the number of items a student had correct on a test specifically designed to measure a student's ability to understand processes of scientific inquiry he used in an investigation conducted in class.*

Process measure score (Part II) refers to the number of items a student had correct on a test specifically designed to measure a student's ability to understand and use the processes of scientific inquiry in a situation quite similar to that performed in class but not identical to it.**

Process measure score (Part III) refers to the number of items a student had correct on a test specifically designed to measure a student's ability to understand and use the processes of scientific inquiry in a science content area not directly related to the area under study.***

---

*See Appendix B for a copy of this test.
**See Appendix B for a copy of this test.
***See Appendix B for a copy of this test.

<u>TOUS score</u> refers to the number of items a student had correct on the <u>Test On Understanding Science - Form eW</u>.

<u>Total process measure score</u> refers to the number of items a student scored correctly on the combined process measure, Parts I, II and III added.

<u>Traditional Curriculum</u> refers to a parallel curriculum constructed along the same "content-lines" as the experimental curriculum. No attempt was made, however, to integrate process and content. The format closely resembled that of the approved texts for Alberta schools for 1967-68 for grades VII and VIII, <u>Science Activities</u> (30).

<u>Hypotheses</u>

The central purpose of the study was to determine if the conscious programming of scientific processes into a science unit would result in measurable gain in a pupil's understanding of these processes as indicated by a number of instruments. Since the main interest was concerned with the attainment of the process-oriented objectives, the major hypotheses deal with this dimension at various levels. It was felt that intensive encounters with processes of scientific inquiry should affect the meaningfulness these processes have for the student. The test scores, therefore, should reflect this aspect as they were an attempt at measuring understanding at various levels.

Hypothesis 1.0 There will be no difference among the treatment groups in the posttest scores on the total process measure, using the pretest scores on the total process measure and the COOP test scores as covariates.

1.1 There will be no difference among the treatment groups in the posttest scores on the process measure (Part I), using the pretest scores on the process measure (Part I) and the COOP test scores as covariates.

1.2 There will be no difference among the treatment groups in the posttest scores on the process measure (Part II), using the pretest scores on the process measure (Part II) and the COOP test scores as covariates.

1.3 There will be no difference among the treatment groups in the posttest scores on the process measure (Part III), using the pretest scores on the process measure (Part III) and the COOP test scores as covariates.

TOUS, according to the theoretical framework upon which it was built, apparently measured objectives quite similar to the process instruments. Its use led to the formulation of the second major hypothesis.

Hypothesis 2.0 There will be no difference among the treatment groups in the posttest scores on the TOUS, using the TOUS pretest scores and the COOP test scores as covariates.

Although the process dimension was the major consideration of this study, the fact that content and process probably cannot be completely separated, pointed to the suitability of administering a more conventional achievement test. A comparison of the results of the respective tests may shed some light on the relationship between the two aspects. In addition, the achievement test was constructed in such

a way that the major portion of the items* were of the comprehension and "higher mental process" type as far as Bloom's taxonomy (7) is concerned. Since almost all of the items on the process measures were also judged** to be of the comprehension and higher mental process types, a definite correlation may exist between the achievement test scores and the process measure scores. Finally, the fact that the process dimension was to be emphasized may have had some effect on the learning of facts and principles; a comparison of the four groups should reveal any tendencies of this nature, if they exist. Hence the third main hypothesis.

Hypothesis 3.0 There will be no difference among the treatment groups with respect to the posttest scores on the achievement test, using the achievement pretest scores and the COOP test scores as covariates.

Limitations

This study suffered from a number of limitations usually incurred by short-term studies of this type. Probably the severest of all was the short, two-month teaching period. For the purpose of determing whether a newly developed curriculum is suitable and workable, this time interval may have been long enough. For evaluation, however, especially where relatively unproven instruments are used, the time interval is almost certainly not long enough. Ideally, studies of this type should be of one or more years duration.

---

*About 70% were of this type.

**The participants of the Project were asked to independently judge each item on the process measure in terms of the category it would fall under in the cognitive domain of Bloom's taxonomy.

Even though it was felt that the teacher variable was a very important one in this study, it was controlled in only a minor fashion. One extra treatment group, made up of classes taught by teachers not participating in the actual development of the experimental curriculum, was added solely for an investigation of the teacher factor. Since one of the basic premises of the approach was the involvement of the teacher in the preparation of the materials, it was anticipated that differences between groups would point to this fact. No other controls such as years of training, teaching experience, teaching load, etc., were thought necessary inasmuch as a number of studies indicate poor correlation between these variables and student achievement.

Other limitations of the study accrued from the limited sampling that was possible. It was anticipated, however, that the statistical design would partially overcome this everpresent limitation encountered when using intact classes. It was felt that a compilation of as much descriptive data as possible would further help to alleviate this problem.

The test instruments which were used were also a very limited sample. The main concern was the most rigorous evaluation possible without the disruption of the ordinary classroom setting. Obviously much more elaborate techniques, such as those suggested by advocates of continuous evaluation, will sooner or later have to be developed.

One obvious limitation was the exclusive concern with the cognitive dimension to the exclusion of domains such as the affective

or psychomotor. It was felt that the concomitant examination of these dimensions would have been too broad and superficial at this stage of development of the project.

III. SIGNIFICANCE OF THE STUDY

More must be known about how students can come to comprehend the totality of science, both substantive and syntactical. Although it is generally conceded that the processes of science are important, very few attempts have been made to incorporate them into the science curriculum, and even fewer attempts at evaluation for their attainment. While a few recent efforts in this regard have been made at the elementary school level, this has been particularly true for secondary school science.

The enunciation of An Inventory of Processes in Scientific Inquiry, stated in such a way that the processes simultaneously act as behavioural objectives, was an important development in this area. Its ability to reflect processes in scientific inquiry as well as to direct the teaching of science, make it a fairly comprehensive instrument.

The difficulty in constructing "higher mental process" questions in science using Bloom's taxonomy (7) of educational objectives has been deplored since its inception. Using the "Inventory" as a framework for constructing "higher mental process" items, however, proves to be a relatively easy task. This aspect further illustrates the versatility of the Inventory .

The effectiveness of using simulated situations in science testing, and the use of audio-visual materials in this area, have similarly not been investigated to a significant extent. Simulation techniques are now being used in universities to prepare students for teaching and for training administrators and appear to show much promise. Perhaps the use of simulated investigations or parts of investigations in testing will prove to be reliable indicators of a student's level of understanding of the process of scientific inquiry.

CHAPTER II

REVIEW OF RELATED LITERATURE

The neglect of the process dimension in science teaching in recent years was certainly not due to a lack of objectives dealing with the processes of science. In 1947 the National Society for the Study of Education (NSSE) stated the following objectives related to the processes in scientific inquiry in their Forty-Sixth Yearbook:

> E. Problem-solving skills, such as ability to:
> 1. Sense a problem.
> 2. Define the problem.
> 3. Study the situation for all facets and clues bearing upon the problem.
> 4. Make the best tentative explanation or hypothesis.
> 5. Select the most likely hypothesis.
> 6. Test the hypothesis by experimental or other means.
> 7. Accept tentatively, or reject, the hypothesis and test other hypotheses.
> 8. Draw conclusions (40).

Unfortunately the general nature in which these objectives were stated did little to suggest how the science teacher could bring about the desired behaviour changes in the students. This situation was partially remedied by Bloom et al (7) in the *Taxonomy of Educational Objectives*. Although the proposed processes or intellectual abilities were only indirectly and generally related to process in science, the emphasis on "higher mental processes" did much to direct science teachers' attention away from mere factual recall in their teaching.

The 1960 Yearbook Committee of the N.S.S.E., concerned with helping the science teacher develop his own objectives, proposed the following guidelines for teaching for the processes of scientific inquiry:

> II. Problem Solving
> Science is a process in which observations and their interpretations are used to develop new concepts, to extend our understanding of the world, to suggest new areas for exploration, and to provide some predictions about the future . . . Methods for solving problems in science are numerous. There is no one scientific method . . . What is done is highly flexible and quite personal . . . The methods of science are something more than measurement, laboratory techniques, and data processing followed by logical deductions. Sometimes they are not very logical, but the search for truth is always present. Presenting problem solving as a series of logically ordered steps is simply a technique to isolate the critical skills and abilities and to give them special emphasis in teaching.
> A process of inquiry involves careful observing, seeking the most reliable data, and then using rational processes to give order to the data to suggest possible conclusions or further research . . .(39).

An even more specific statement of objectives or expected outcomes related to the process dimension was that developed by Obourn. He analyzed the problem solving abilities which can be developed through science teaching into some eighty more or less specific skills. The seven general categories he proposed were:

A. Formulates significant problems.
B. Analyzes problems.
C. Obtains information regarding a problem from a variety of sources.
D. Organizes the data obtained.
E. Interprets organized data.
F. Tests the hypotheses.
G. Formulates a conclusion (43)

It is hard to imagine why the teaching of science did not change drastically in order to embrace these numerous objectives. Even the "new science curricula", although stating objectives very similar to the ones cited above, only partially succeeded as far as process objectives were concerned. What appeared to be urgently needed were theoretical frameworks from which would flow the arguments in support of the process dimension in science teaching. More importantly, though, implications for the classroom had to be spelled out in detail.

The general dissatisfaction with science teaching, then, and the incomplete approaches taken by the "new science curricula" led many science educators to re-examine the fundamentals upon which these courses, and science education in general, were built. One of the most publicized analyses and subsequent expositions on this topic came from Joseph J. Schwab in _The Teaching of Science as Enquiry_ (54), the Inglis lecture for 1961. Using the very nature of science for support, Schwab postulated that the structure of any given discipline, such as science, consists of the _substantive_ aspects (ideas, conceptual schemes and modes of organization) and the underlying syntactical structure (kinds of evidences required, how far the data are obtainable, what set of second-best substitutes are employed, et cetera). In the vernacular, these two dimensions are referred to as the _product_ (content) and the _process_ of inquiry, respectively.

The implications became clear as Schwab elaborated:

> A completely enquiring classroom requires teaching and learning skills which are not the common habits of our schools. Its aim is not only the clarification and inculcation of a body of knowledge but the encouragement and guidance of a process of discovery on the part of the student.
>
> For the student, this means relinquishment of habits of passivity, docile learning, and dependence on teacher and textbook, in favor of an active learning in which lecture and textbook are challenged. The lecture and textbook cease to be authoritative sources of information to be learned and become materials to be dissected, analyzed. For, in one form or another, the materials of such a classroom are not statements of truth but reports of enquiry. Hence the student's attention is not on something said but on something done. The oral and written material presented to him still, inevitably, are sayings. But the student's attention is not on the statements as statements - words and assertations to be learned - but on what the words and assertations are about: the thoughts and actions of a scientist which have gone into the making of a piece of scientific research (56).

It is pointed out that the emphasis in teaching must still be on content but primarily on those significant concepts which define the substantive structure of a subject. However in addition to the matter of selection, the organization of content has a key role in learning. According to Bruner, "subjects have a fundamental structure that is basic to effective learning and must therefore be central in teaching (9)". Bruner further described the impact of a discipline on teaching and learning as follows:

> Intellectual activity anywhere is the same, whether at the frontier or knowledge or in the third grade classroom. What a scientist does at his desk, or in the laboratory, what a critic does in reading a poem are of the same order as what anyone else does when he is engaged in like activities - if he is to achieve understanding. The difference is in degree not in kind. The schoolboy learning physics is a physicist,' and it is easier for him to learn physics behaving like a physicist than doing something else (9).

Hilda Taba (62), in her four levels of knowledge, presented ideas very similar to those of Schwab and Bruner. She suggested four levels of knowledge: (a) specific facts and processes, (b) basic ideas, (c) concepts, (d) thought systems. It was the fourth level that suggested:

> . . . that the academic disciplines represent thought systems and methods of inquiry as well as compendiums of knowledge. These thought systems are composed of propositions and concepts which direct the flow of inquiry and thought. Presumably, each discipline represented by a school subject is organized around some such system of interlocking principles, concepts and definitions. These systems direct the questions asked, the kinds of answers sought, and the methods by which they are sought. (45).

The National Science Teacher's Association (NSTA) in a recent publication (38) provided the following general guidelines for science teaching:

> . . .2. A heavy emphasis should be placed on the nature of science or the process by which new knowledge is obtained. 3. Instruction should be planned to develop understanding of basic ideas of science concomitant with an appreciation of the methods of science; these two aspects should not be treated independently (38).

The Process - Approach in Science Teaching

How have these rather general theoretical principles and positions been applied in the classroom? One of the first research projects undertaken was that by J. Richard Suchman (61). He attempted to put Schwab's ideas on the teaching of science as inquiry into operation by examining a limited aspect of science teaching - the

verbal interaction between the students and the teacher while discussing and investigating posed problems. Using aspects such as fluency, the number of questions that a student posed, and the frequency with which certain types of functional questions were used, Suchman was able to assess a student's growth in ability to inquire. This teaching strategy, however, had obvious limitations since mainly verbal behaviour was analyzed and investigated. As far as the nature of science is concerned,then, this method of teaching science allowed for the learning of some of the processes of scientific inquiry. The emphasis, however, was on being able to use these processes at the verbal level rather than in the laboratory situation.

Another promising attempt at teaching for the processes of scientific inquiry was that developed by Klopfer and Cooley (35) using their now famous "case history" approach. They conjectured that:

> Probably the principal reason for the disparity between the frequent objectives concerning student understanding of science and scientists and the apparent shortcoming of their attainment in the minds of students is that very few instructional procedures specifically designed for this purpose have been proposed (36).

Being influenced by Conant who used case histories at the university level (15), Klopfer believed that one way to get at the many intangible objectives repeatedly stated in science curricula would be for the student to investigate the history of science. He developed a number of case histories representative of the physical and biological sciences which, if used in the classroom, would enhance a student's understandings about the processes of scientific inquiry, the scientific enterprise,

and the characteristics of scientists. Even though original and similar experiments were to be duplicated wherever possible, the approach was still rather verbal and isolated in nature as no attempt was made to build a general framework into which these individual case studies could be fitted.

It was only a matter of time until someone or some organization undertook to build a completely process-centered curriculum in science. The American Association for the Advancement of Science (AAAS) under the research direction of Robert M. Gagné did just that by proposing a science curriculum which was process-based and hinged heavily on the psychology of learning developed by Gagné (24). The course of studies they proposed was called Science - A Process Approach (1).

Gagné outlined the following rationale for the "process approach":

> This approach seeks a middle ground between the extremes I have mentioned the "content" approach and the "creativity" view. . . Specifically, it rejects the "content approach" idea of learning highly specific facts or principles of any particular science or set of sciences. It substitutes the notion of having children learn generalizable process skills which are behaviourally specific, but which carry the promise of broad transferability across many subject matters. . .. The point of view is that if transferable intellectual processes are to be developed in the child for application to continued learning in sciences, these intellectual skills must be separately identified, and learned, and otherwise nurtured in a highly systematic manner (25).

One of the strengths of this curriculum was that it was designed to be introduced at the kindergarten level as opposed to other studies

which directed their attention to junior high or high school. In this way the project attempted to avoid the "set" that students, taught in conventional ways, develop against active participation, individual work and self-directed learning in general. As Ivany claims:

> A novel teaching situation can prove difficult for a class of students merely because of the confrontation with novelty. Because of a set established by years of experience in almost totally expository classroom environment, a presentation of a lesson in the "hypothetical" mode would be perhaps what Thelen refers to as a violation of expectations. Such sets, or dominant tendencies to behave in a particular manner, could well be a problem in any school experiment (34).

Gagné's Psychological Bases of Science - A Process Approach outlined the basic premises of the "process-approach" in the following manner:

1. The scientists' behaviours in pursuing science constitute a highly complex set of intellectual activities which are, however, analyzable into simpler activities.
2. These intellectual activities (processes) are, as most scientists would agree, highly generalizable across scientific disciplines. . .
3. These intellectual activities of scientists may be learned, and it is reasonable to begin with the simplest ones and build the more complex activities out of them, since this seems to be in fact the way they are organized.
4. Accordingly, one can construct a reasonable sequence of instruction which aims to have children acquire process skills, beginning with the simplest kinds of observation, and building progressively through classifying, measuring, communicating, quantifying, organizing through space and time, to the making of inferences and prediction. As further building occurs, one finds it possible for students to learn how to make operational definitions, how to formulate testable hypotheses, how to carry out experiments,

and how to interpret data from experiments. . .

5. At the end of such instruction, the student will not necessarily know anything which can be identified as physics, or chemistry, or biology, or geology. What will he know, then? Perhaps something like this: A scientist should be able to tell this student what he (the scientist) is studying and the techniques he is using, and what he has found, in a relatively brief fashion, and to have the student display a rather profound understanding of it immediately. . . (26).

Although the program was severely criticized on a number of issues by people like Atkin (3) and Ausubel (4), it has a number of strengths which should not be overlooked. First, the program had a thorough rationale obviously influenced by Gagné's learning theory. Second, a painstaking effort was made not only to define all objectives in terms of observable behaviours, but also to break down each process into a large number of basic skills whose inter-relationships were presented graphically. Third, evaluation was an integral part of the program since its inception. Fourth, since the program was activity-oriented, emphasis was on active participation on the part of the student with ample opportunities for laboratory-type experiences.

In order to further appraise the efforts of this organization critically, an exposition by Parker and Rubin was used (45). In a recent synthesis, they saw the following requirements for a process-based curriculum:

> The requirements posed by a process-based curriculum deal primarily with the identification of worthwhile processes to which students should be exposed, the design of instructional strategies that make effective use of the processes, and the realignment of subject matter so that it complements the instructional strategies. Each of the requirements must be the product of pointed research and field testing, each must have roots in theory and function, and each must reflect the character of the discipline, or disciplines, under which it operates (46).

As far as the first requirement was concerned, the AAAS course of studies did identify a number of "knowledges" (1) which were felt to be fundamental to understanding and engaging in the processes of science. The exact source of these basic knowledges or intellectual activities was not given; they seem to have been compiled in a number of discussions about the behaviours that scientists display while pursuing science. The exact interrelationships among these processes were not defined; only a psychological hierarchy from simple to complex skills and understandings was envisioned. Perhaps the greatest shortcoming of the program, however, was due to its failure to "reflect the true character of science" (46) - the last requirement cited by Parker and Rubin. It seemed that the substantive aspect of science was left as much to chance as the syntax was left in traditional courses. This weakness was partially revealed by Gagné's own admission that he was not certain of the answer to the question: "What does this mean the student is ready for in terms of additional science instruction?" (27). It is difficult to concur with the view that "such a student should be able to learn any given science, in terms of its theoretical structure, in about half the time that it would otherwise require" (21) - at least

not in terms of Schwab's conception of the structure of a discipline. That is, if Schwab's ideas on the two dimensions of the structure of science are accepted, then it is improbable that these students will come to comprehend the substantive structure of science.

The instructional strategy of having children acquire basic process skills such as observation, classification, prediction and inference which are to eventually allow him to hypothesize, experiment and build models seems rather difficult to defend unless it is postulated that children of elementary school age are incapable of carrying out these basic skills in the more general framework of experimental investigations. It should be possible to learn these processes in their proper context of actual experimentation in order for the student to see interrelationships and in order for him to become familiar with the totality of the syntax of science.

During the last two years, the Edmonton Junior High School Science Project, under the direction M. A. Nay, has been developing a new approach for teaching for the processes of scientific inquiry. Dissatisfied with the partial approaches outlined above, this group set out to determine a theoretical framework upon which a science curriculum recognizing the importance of the process dimension could be built. Using summer institutes for junior high school students, classes taught by the inquiry method were carefully observed in order to isolate processes and dimensions which seemed to affect this type of teaching. On the basis of this experience and the adoption of

some of the ideas discussed earlier, the development of a theoretical framework was undertaken.

This theoretical framework would hopefully (1) guide the programming of the proposed unit, (2) indicate special teaching strategies, and (3) serve as a framework for evaluation. In a sense it is somewhat ecletic in that it is an amalgam of a number of theoretical positions and recent developments in science education. First, Schwab's (53) theory of the structure of the disciplines is integral in that subject matter is considered primary, dictating the problems and concomitant processes to be investigated. That is, emphasis is on Bruner's (9) significant and pervasive concepts which define the substantive aspect of the science under consideration. Second, the contribution of the Commission on Science Education's *Science - A Process Approach* (1) was acknowledged by using their process-approach as a basic strategy. The generality of the identified processes was recognized with the realization that they become more specific when the context of their use is specified. In addition, the processes are stated behaviorally so as to facilitate evaluation for growth in *knowledge* and competence in the *use* of these processes. Third, it was considered necessary that the processes be consciously integrated into the program by developing special teaching strategies. In Schwab's words:

> For the teacher, the transformation of the conventional classroom into a completely enquiring one also demands new skills and habits. A student does not learn to

> "learn for himself" merely by being told to do so. Still less can he discover for himself what sorts of parts exist in a scientific enquiry, what their roles and connections are, and so on. He cannot be expected, in short, to know automatically, what to look for in a report of scientific inquiry, what questions to ask of the material he is reading. On the contrary, this is the first and major responsibility of the teacher.
>
> In the dogmatic classroom, the role of the teacher was to explain what the book left unclear and to test the student's grasp of what he was told. Now, his role is to teach the student how to learn. His responsibility is to impart to the student an art, a skill, by means of which the student can teach himself. This art consists of knowing what questions to ask of a report of enquiry, when to ask them, and where to find the answers. This kind of skill is learned by doing, by exercise, and is taught by guiding the doing (57).

The arrangement or structuring of these processes was probably a result of the popularity of the "inventory-approach" in analyzing "the way teaching is".* As is evident from the framework, then, the heart of the process approach being developed is _An Inventory of Processes in Scientific Inquiry_.**

In this "Inventory" a number of identifiable processes in scientific inquiry are listed and organized. These are processes which are generally observed when scientists are at work. However, the particular character of each scientist _and_ the specific aspects of each investigation are recognized and in no way discounted. But, in order to communicate and to teach about the scientific enterprise,

---

*A large number of observation schemes analyzing classroom interaction have adopted the "inventory-approach". The ASCD's _The Way Teaching Is_ (2) describes some of these schemes.

**See Appendix A for a copy of the "Inventory".

the common rather than the specific acts and processes needed to be identified first.

The Inventory also relates very consistently to various levels of scientific inquiry which are determined by the amount of guidance a pupil receives in any particular investigation. Schwab explained this idea in the following way:

> Three different levels of openness and permissiveness are available for such invitations to laboratory enquiry. At the simplest level, the manual can pose problems and describe ways and means by which the student can discover relations he does not already know from his books. At a second level, problems are posed by the manual but methods as well as answers are left open. At a third level, problems, as well as answer and method, are left open: the student is confronted with the raw phenomenon - let it be even as apparently simple a thing as a pendulum. He pushes and pulls, alters first one and then another of its aspects, begins to discern a problem to be solved, then moves toward its solution (55).

That is, if the whole investigation was structured for the pupil, the approach essentially was a "cook-book" one; on the other hand, as more and more guidance was removed, the approach became one of problem solving and finally problem making.

For the purpose of structuring the process strategy, the seventeen gross processes which were identified were grouped under five major headings:

I. PREPARATION
II. COLLECTION OF DATA
III. PROCESSING OF DATA
IV. CONCEPTUALIZATION OF DATA
V. OPEN-ENDEDNESS*

---

*For a more detailed list and explanation of the various processes see Appendix A.

Nevertheless, the identified processes can also be divided into sub-categories which would necessarily become more specific in terms of a given problem being studied. Grouping the processes under five major headings is analogous, in some respects, to the "formal" way of presenting science experiments in laboratory manuals. Hence, a formalism for teaching is suggested which is reflective of Dewey's (17) five steps of problem solving and at the same time recognizes the tradition and realities of science teaching.

It is to be emphasized at this point that although the processes are numbered from one to seventeen, no rigid sequence or order is implied. That is, it is conceded that not all investigations proceed in an orderly fashion from step one to seventeen. Moreover, the fact that this Inventory was made as comprehensive as possible does not imply that every investigation included all steps from one to seventeen, but rather that whatever steps were found in an investigation could also be detected in the Inventory.

This Inventory was used in building a science curriculum on the topic of "Matter and Energy" in which both the syntactical and the substantive nature of science were recognized. Content was first structured into manageable units upon each of which the Inventory was superimposed according to the syntax expounded by Schwab. In this way, not only is the Inventory a fairly valid representation of what scientists do when they go about doing their research, but it is also a useful formalism for laboratory-oriented science teaching.

The basic teaching unit in this experimental program on "Matter and Energy" was the "Investigation".* After brief preliminary sessions in which the student was exposed to the processes, he was given ample opportunity to use them consciously in learning science. As many scientific processes as feasible were carefully programmed into each investigation. The criterion of programming was the conscious effort to make the student more and more self-reliant in carrying out successive investigations. That is, a progression from step-wise guidance at first to self-directed activity on the part of the student was envisioned. It was felt that only by making the student participate in learning science in this fashion would he assimilate a true picture of the character of the scientific enterprise.

How did this experimental curriculum, on "Matter and Energy", meet the requirements for a process-based curriculum as seen by Parker and Rubin? First, worthwhile processes to which students should be exposed were identified. Secondly, instructional strategies were specifically designed to teach the identified processes. And finally, subject matter was realigned to complement these instructional strategies wherever possible.

<u>Evaluation in Science Teaching</u>

As already indicated, the "new science curricula" of the 1950's and early 1960's, although stating objectives related to process in

---

*See Appendix A for a sample investigation.

science, generally neglected to test their pupils to determine whether these objectives had actually been attained. An examination of an annual compilation of curriculum developments in science and mathematics (14) indicated that the majority of new studies neglected to develop any evaluation instruments. Those who did, made up their own "content" tests usually based on the theoretical framework provided by Bloom's Taxonomy (7).

It is not surprising then, that efforts to measure the effect of trying to teach by inquiry, the investigative approach, and problem solving have revealed no significant differences among treatment groups when traditional tests were used to measure growth relative to new objectives. It seemed that:

> Science educators must face up to the fact of the failure of their programs to teach in the hoped-for manner or assume that significant gains were made but that the testing instruments used were not capable of measuring that gain (35).

Many researchers in this area hold to the second point of view.

Klopfer and Cooley (35) recognized the need for "new" types of tests to accompany novel efforts in the teaching for the nature of science and scientific inquiry such as their case history approach. Accordingly they identified a number of objectives, dealing with process, which were more or less common or recurring from one case history to the next. Provided with funds from the United States Office of Education, they set about constructing and validating the _Test On Understanding Science_ (TOUS) - Form W. Although more general, this test, being a measure of a student's understanding about the processes of scientific inquiry, the scientific enterprise and the characteristics

of a scientist, measured objectives quite similar to many found in the Inventory. Unfortunately however, the original TOUS Form W was designed for high school whereas this study was concerned with grades seven and eight. The more recent Form eW, which was developed for the Elementary School Science Project (ESSP) and the School Science Curriculum Project (SSCP) to measure similar objectives but at the upper elementary and junior high school levels was found very helpful in this regard.

Members of the evaluation group of Harvard Project Physics (28) have recognized the need for multi-dimensional testing in order to determine what students really achieved when they were taught by this particular Project. The spectrum of measures they chose was:

A. Measures of change
   1. Physics achievement
   2. Interest in science
   3. TOUS
   4. Attitudes toward various aspects of physics such as doing laboratory experiments, physicists, and "myself as a physics student"
   5. Knowledge of the processes of science (Welch Science Process Inventory) (65).

Although not all of these are measures related to cognitive aspects of content and process in science, the process dimension has been given recognition through the use of the TOUS and the Welch Science Process Inventory.*

In a similar effort, the Portland Public School system, being an official tryout centre for a large number of "new science curricula", saw the need for testing a wider range of objectives than those suggested by conventional achievement tests. Accordingly they developed the Portland Science Test (49) which attempted to measure gains in

*At the time of the writing of this thesis this particular test was in press otherwise it too might have been used.

students' understandings of both the product and the processes of science. On this test, the student was asked to respond to process and product-type questions related to pictorial stimuli with which he was presented.

Another attempt at an evaluation of the process dimension was made by the Biological Science Curriculum Study (BSCS) (5). This curriculum study, with publishing assistance from the Psychological Corporation, pioneered in this area with a test called the Processes of Science Test (6). The authors claim that it is:

> a measure of student understanding of general scientific principles, the nature of scientific inquiry, and scientific reasoning. . . Objective questions test the student's ability to evaluate criteria for accepting or rejecting hypotheses and to appraise such aspects of experimental design as the need for controls, repeatability, adequate sampling, and careful measurement (51).

This test made use of a simulation technique in that various aspects of scientific investigations such as numerical data, graphs, paragraphs of procedure, et cetera were presented to the student in order to simulate an actual investigation. A number of multiple-choice questions followed, usually four or five, which were based directly on the simulated investigation presented in the introductory diagrams or paragraphs to each particular set of questions.

It was felt that this technique could be very useful in testing for process in science teaching. At the same time one serious limitation was recognized - the fact that a rather extensive amount of verbiage had to be used in order to simulate a scientific experiment. The Portland Science Test cited earlier partially overcame this difficulty by

restricting the stimuli presented to the student to diagrams and pictures. In this way they expected to meet the goal of evolving a test instrument which:

1. places a minimum reliance on reading skills
2. evaluates knowledge of the ways in which scientists learn, as well as what scientists have learned
3. reveal the emphasis in the Portland Science program on investigative patterns of learning (50).

The authors of the test, in using sketches and pictorial stimuli, gave the following argument in support of their stimulus mode:

> Teachers regularly observe the student interacting with a world of natural events and learning to represent this experience through the use of symbols and signs. However, it is obvious that students do not soon learn to capture the world with words. Therefore it was necessary to seek a test stimulus mode that would stimulate the student to reveal, in the test situation, his functional level of understanding of natural events, not just his ability to handle abstract symbols (50).

In selecting such a stimulus mode, the committee anticipated the work of Schalock, Beaird and Simmons (52) in the area of teacher education. These researchers postulated a "Continuum of Test Stimulus and Response Complexity" (Figure 1).

They hypothesized that the predictive power of the tests would be largely determined by the relative position of the stimuli on the continuum (52). Thus, the MTAI test was seen as the weakest predictor and the Simulation as the strongest.

SITUATIONAL RESPONSE TEST

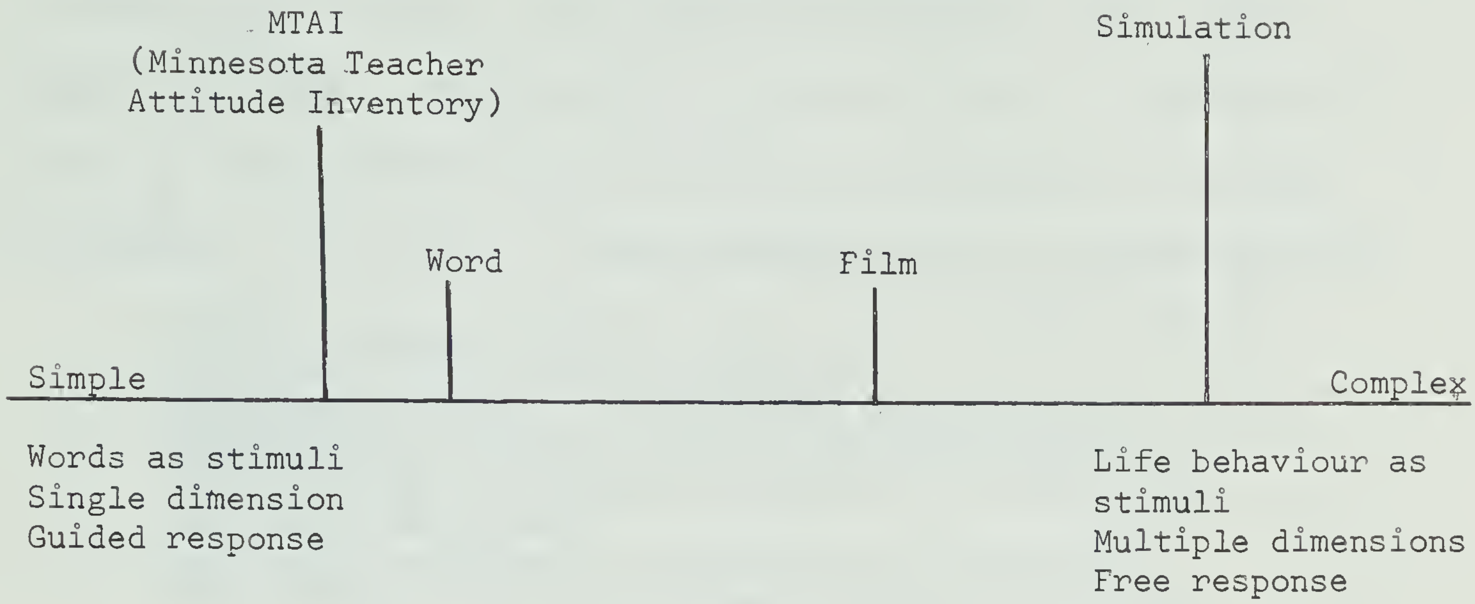


Figure 1. Continuum of Test Stimulus and Response Complexity.

Arguing thus, the Portland Science Test would fall somewhere between the Word test and the Film test.

Going one step beyond the <u>Portland Science Test</u> in a manner similar to but not identical with a more recent format developed by the Portland Public School System, this study used single concept loops to present stimuli to the student in a test situation. It was felt that "this stimulus mode more nearly approaches the physical configuration of that evoking the real life behaviour of students in the science classroom laboratory" (52). Although the film loop format allowed the simulation of the visual experiences of the student in the science classroom laboratory, the response mode was still the familiar multiple choice, verbal one.

Using <u>An Inventory of Processes in Scientific Inquiry</u> as a framework for choosing questions to which the student responded, a pilot study was carried out to construct and validate a nine-item

test* for each film loop. But what situations should have been used in these filmed episodes? Some recent theorizing** in this regard helped to provide a rationale for the choice of films. A hierarchy of levels of "meaningfulness" at which these processes may be known was hypothesized. This scale of meaningfulness could be represented thus:

1. Intuitive grasp referred to the use of a process without being consciously aware of using it.
2. Knowledge referred to knowing what is meant by a process.
3. Comprehension referred to displaying knowledge and understanding of a process by using it in a situation previously encountered.
4. Application referred to displaying an understanding of a process, at a higher level than above, by using it with content that is similar to but not identical with that already encountered - a kind of specific transfer in Bruner's terms.
5. Transfer (Ability to learn) referred to displaying knowledge and understanding of a process in any situation with which a student is confronted, science or otherwise - a sort of a general transfer.***

Incorporating the above rationale, the understanding of the various processes was tested at three levels: comprehension, application and transfer. That is, of the nine film loops chosen, three fell into each of the three categories. Three film loops were chosen or filmed

---

*Three sets of nine items were chosen because the showing of three films was suited to a 45 minute class period.

**The writer is indebted to M. A. Nay and Robert Crocker for their assistance in helping to crystallize this particular rationale of a hierarchy of "meaningfulness".

***The term "ability to learn" was used by Leo Nedelsky (42) in referring to the highest level of cognitive objectives.

representing investigations actually performed in class (comprehension). Three additional film loops were chosen to represent situations involving similar content to that covered in class, but the investigations were not actually performed (application). The last three film loops dealt with physical science content only incidentally related to the content covered in class. It was anticipated that an investigation of the results obtained through the use of these tests would yield some interesting information on the adequacy and relevance of the hierarchy indicated above; and perhaps more important, shed some light on the little-understood relationship between process and content in science.

Continuing with the survey of attempts of finding suitable means of evaluating pupil growth along the process dimension, the work of the AAAS in *Science - A Process Approach* could not be overlooked. Perhaps their major contribution to this area was their recognition that *behavioural* objectives were a necessity before any evaluation was attempted. Accordingly, each exercise was preceded by a specific set of objectives whose degree of attainment was to be specifically and rigorously determined. The nature of the objectives was such that achievement could only be judged by observing an individual's performance. This may have been possible for simple skills, but it became exceedingly more difficult, time-consuming and impractical as the skills became more complex. Not only does the ordinary elementary school teacher not have the time for individual testing, but judging

the degree of attainment of complex skills becomes an exceedingly ambiguous and arbitrary task. It should be emphasized nevertheless, that the rationale of evaluating progress or growth in small steps was a psychologically sound one, even if very difficult in practice.

Recent developments in computer applications to education,* have provided a ray of hope in this regard. Underlying computer-assisted-instruction (CAI) is the rationale that the computer keep a continuous record of a pupil's performance in order to use the student's feedback in deciding the psychologically and logically soundest path for him to follow in future learning. The only real obstacle holding back progress is the development of natural languages and the accompanying "software" which will perform the necessary functions such as record-keeping, providing instructional sequences and proper handling and processing of pupil feedback.

The process approach developed in this investigation lends itself very well to a continuous analysis such as that suggested by CAI. The student can keep a very detailed notebook containing his ideas and performance on the processes he encounters. As is the case with CAI, a scheme or model, based on pupil performance, can be devised which will

---

*The Portland Public School System, under the direction of Dr. Donald Stotler, has been very active in this area. They have been trying to adapt the AAAS approach to CAI by working out the CYBEX and TAG systems (31, 32, 33, 47, 57, 58, 59). Also, in the Intermediate Science Curriculum Study (ISCS), under the direction of Dr. E. Burkman at the Florida State University, computer applications in the development of a new curriculum has been examined.

allow the determination of a student's progress in the process dimension. This means of evaluation was not undertaken in this study as there are no appropriate models available as yet. It is felt that setting up a model of this type could be the subject of a subsequent dissertation. The students notebooks were scrutinized, however, by a few participating teachers who were seeking clues to strategies which might be used in "process" teaching.

CHAPTER III

METHODS AND MATERIALS

At the time this study was conceived, the Edmonton Junior High School Science Project had already been set up. The first objective was to develop a process-oriented curriculum on "Matter and Energy" on the basis of the Inventory. This program will hereafter be called the "Experimental" program. The second objective of the project was to teach this Experimental program to students of the science teachers involved in the project. The writer joined the group when he was offered the opportunity to take charge of the evaluation aspect of the project.

The research group, then, was primarily interested in whether the teaching for processes in scientific inquiry in the ordinary classroom setting was possible. The author's main interest was whether pupil growth in this dimension could be evaluated in some feasible fashion. Hence at his suggestion, a number of controls were imposed such as the number and types of classes selected, the number of treatments, selection of pre and posttests, types of schools selected, et cetera, in order that at least a quasi-experimental design could be approximated.

I. THE SAMPLE POPULATION

The classrooms of the teachers participating in the research group, then, made up one experimental treatment group. The other three treatment groups were composed of classes found in three different schools which were roughly equated in terms of socioeconomic status by selecting suburban junior high schools. The schools and not the classes, however, were assigned to treatment groups in order to avoid contamination of the teaching strategies from one teacher to another in any particular school.

Sample

Grade seven and eight classes were selected from the participating schools* and these schools were then randomly assigned to the various treatment groups. The classes in Treatment A were, of course, not randomly assigned. Table I shows the distribution of students by sex, class, grade and experimental group. The fact that intact classes were used caused totals in the treatment groups to vary somewhat. Grade seven and eight pupils were chosen because of their relative availability for experimentation and because of their unfamiliarity with the chosen subject matter "Matter and Energy", which is usually a grade nine topic in Alberta schools.

---

*Essentially the classes were selected on the recommendation of R. Melnychuk, Science Supervisor of the Edmonton Public School Board. He recommended them on the basis of the willingness of the teachers to cooperate and according to their availability for experimentation. A large number of grade seven and eight classes were not available as they were on experimental programs for the Provincial Department of Education.

TABLE I

DISTRIBUTION OF STUDENTS

Group	A						B					C							D				
Class	11	12	13	14	15	16	21	22	23	24	25	31	32	33	34	35	36	37	41	42	43	44	45
Grade	8	7	8	7	7	7	8	8	8	8	8	8	8	8	7	7	7	7	7	7	7	7	8
Boys N	15	14	13	23	27	13	16	10	11	16	11	15	13	15	14	21	24	21	13	14	13	17	19
Girls N	18	14	20	--	--	15	14	18	15	13	17	14	13	11	18	11	11	12	15	8	18	11	14
Class N	33	28	33	23	27	28	30	28	26	29	28	29	26	26	32	32	35	33	28	22	31	28	33
Group N	172						141					213							142				
Total N	668																						

II. TREATMENT PROCEDURES

<u>Description of the Four Treatments</u>

An attempt was made to teach the "Matter and Energy" content to three of the four treatment groups. The groups differed essentially in the emphasis placed on the teaching for and learning of the processes of scientific inquiry.

<u>Treatment A.</u> This first experimental group was instructed by teachers who, as members of the Edmonton Junior High School Science Project, helped to develop the Experimental program on "Matter and Energy". As such, it was assumed that these teachers not only knew the processes to be identified and emphasized, but also understood them to a greater degree than the teachers not participating. Materials were prepared for the teachers and the students in order to control differences among teachers to some extent. A conscientious effort was made in the prepared materials to integrate, as much as possible, the content and the associated processes deemed useful.

<u>Treatment B.</u> This group also was taught the Experimental program on "Matter and Energy". The basic difference between the first and second group was that the teachers in this treatment group were not "brain-washed" in the origin and substance of the Inventory or its applicability and usefulness in science teaching. No intensive effort was made to explain this Inventory to these teachers nor to influence their teaching in any manner. They did, however, receive the same prepared materials to teach and distribute to their students as

treatment group A. Thus this mode provided some indication of the necessity, importance of and the amount of inservice work required before and during the teaching of these materials, which however was not provided.

Treatment C. This group was taught the same content as the first two experimental groups, but no attempt was made whatsoever to incorporate the processes of science as identified in the Inventory . A so-called "Traditional" program on "Matter and Energy" was produced. The teachers were given prepared textual material modelled after the approved texts for grades seven and eight in Alberta Schools: Science Activities (30). No special instructions were given to the teachers except to continue teaching in the same manner as they had prior to the experimental program.

Treatment D. This treatment group provided a control for the other treatment groups. The regular yearly teaching schedule of this group of classes was in no way affected, except for testing, as no special curriculum was prepared. It was anticipated that factors such as maturation, effects of pretests and testing in general could be controlled to some extent with a group of this type. The content covered in the classes for the time interval between tests was only incidentally, if at all, related to the content of the tests which were administered.

Testing Procedures

At the beginning of the experimental period in each class, students were given the Cooperative General Science Test - Form B,

Before any further testing, each class was given a two-period instruction dealing with the processes of science as listed in the Inventory . This instruction was necessitated by the possibility of obtaining variations in results due to a lack of knowledge of the vocabulary used on the pretests, especially on the process measures. Then each treatment group in turn and in random order, was pretested on the five main instruments used in the study: the TOUS, the conventional achievement test and the three process measures. At the end of the experimental period, roughly of two and one-half months duration, all of the five major instruments were readministered to measure changes. Each of the instruments is discussed below.

III. THE FILMED EPISODES

A number of considerations was important in choosing topics for the nine film loops. An attempt was made, first of all, to sample adequately the content of the unit under study. A great amount of leeway was possible as a result. Unfortunately, however, the availability of film loops in the "Matter and Energy" content area was still quite limited and only seven film loops were found* which dealt with the proper subject matter. Two additional film loops were prepared by the writer** to make up the required number.*** Another feature of selection was that the situations presented

---

*Four film loops from Suchman's Inquiry Development Program (62) were selected along with two loops kindly submitted by the Edmonton Public School Board and one loop from the Edmonton Separate School Board.

**The staff of the A-V Media Center of the University of Alberta helped in the preparation of these last two films.

***Appendix A lists the films by title and gives the sources from which the film loops were obtained.

on the films would be similar to those the students encountered in the learning situation. That is, simple equipment and unsophisticated techniques were used in a classroom-laboratory setting. Because the topics covered in the film loops are not usually part of the grade seven or eight science curriculum, it was anticipated that the content of the film loops would be novel. The recency of development of film loops in general further helped to reduce the possibility of students having encountered these films previously.

All films were tried for suitability during a pilot study and it was during this phase that mechanical details of administration were worked out. As a result of the pilot study experience, it was decided to administer the tests which accompanied the film loops in the following manner: the class was called to order and the matching test booklets and answer sheets were handed out; the class was told to observe the sequence of events on the film loop carefully; students were then allowed to read the stems of the nine accompanying questions; the film loop was reshown; and finally the pupils were given eight to ten minutes to answer the nine questions. The same procedure was followed for the other two film loops thereby allowing the administration of three film loops during one forty-five minute period. The other two sets of three film loops required two additional testing periods.

IV. TESTING INSTRUMENTS

Cooperative Science Test

Form B of the general science test was specifically designed for junior high school students. It consists of sixty multiple choice questions to be completed in a forty minute testing period. This test was chosen as an indicator of a student's level of achievement in general science because its statistics are well documented (19) and an examination of the types of items and their content showed the test to be valid for administration at this level. In addition it correlates rather highly with the School College and Ability Test (SCAT) (19). Additional arguments for validity stem from its intensive rebuilding from earlier cooperative tests and the fact that many well-known scientists participated in the construction of the final form (20). Furthermore, the indicated breakdown of subject matter upon which individual items are based and the relative emphasis of objectives are thought to be suitable indicators of a student's level of achievement in science.

Reliability estimates for a population of 450 students from each of grades seven and eight are given as 0.90 (21) and are of the internal consistency type, computed using the Kuder-Richardson Formula 20. This estimate as published in the Cooperative Science Tests Handbook is reproduced in Table II.

TABLE II

RELIABILITY ESTIMATES OF TESTS*

Test	Reliability	Highest Correlations with other Variables
Cooperative Science Test	.90	.70 (Post achievement) .69 (Total Process-Post)
Achievement Test (pretest)	.73	.66 (Post achievement) .65 (COOP Test)
TOUS (pretest)	.83	.64 (Post-TOUS) .54 (COOP Test)
Process Measures (pretests)		
Part I	.79	.76 (Total Process-Pre) .63 (Post - Part I)
Part II	.74	.73 (Total Process-Pre) .67 (Post-Part II)
Part III	.78	.71 (Total Process-Pre) .52 (Achievement Pre)

*All reliability coefficients were calculated using the Kuder-Richardson Formula 20.

Achievement Test

The major purpose of this test was to see what effect, if any, the special emphasis on processes in the experimental groups had on a pupil's performance on this measure of conventional achievement. Fortunately, a bank of multiple choice items*, based on Bloom's Taxonomy (7) and dealing with the content of the study, already existed. Fifty items were chosen or specially constructed and administered to 136 students in a pilot study. The purpose of the study was two-fold: revision of items based on an item analysis and the estimation of a coefficient of reliability. This coefficient was found to be 0.81 using the Kuder-Richardson Formula 20. A subsequent coefficient of reliability, using the same formula, was calculated using the pretest data. This coefficient is reported in Table II.

*In the school year, 1966-67, a group of junior high school science teachers of the Edmonton Public School Board set up a bank of multiple choice items for the grade nine content of which "Matter and Energy" is a part.

Test On Understanding Science (TOUS)*

As indicated previously, this test was based on objectives not unrelated to those outlined in An Inventory of Processes in Scientific Inquiry. Although items dealing with characteristics of scientists were also included in this test, it was felt that this test was one of the few existing ones which concerned itself with process in science at the desired grade level. As such its correlation with the process measures may be an indication of the validity of the process measures. The TOUS, Form eW, has undergone rigorous testing and the reliability estimates, as given by the Kuder-Richardson Formula 21, are 0.70 for grades seven and eight.** Using pretest data, a coefficient of reliability of 0.83 was obtained based on the Kuder-Richardson Formula 20. This coefficient is similarly reported in Table II.

Process Measures

These instruments were specifically designed and constructed to measure growth along the process dimension since no other suitable test instruments were available. Their validity is based of the fact that they reflect the Inventory of Processes in Scientific Inquiry which in turn reflects the syntactical nature of science. The members of the

---

*The writer is indebted to Dr. Peter B. Shoresman, Associate Professor of Science Education at the University of Illinois, for his permission to use the TOUS - Form eW.

**This and other coefficients of reliability for TOUS, Form eW, were obtained from a personal communication from the Elementary School Science Project, University of Illinois.

research group, being thoroughly familiar with this Inventory, classified each constructed item to determine if it was a measure of any of the seventeen processes in the Inventory. Items revealing major disagreement were discarded and the final tests consisted of nine film loops accompanied by nine multiple choice questions on each. Another attempt at ensuring validity was the administration of these tests to classes known to have participated in process-oriented teaching for some time and others known to have been taught in a more or less traditional fashion during their junior high school years. Table III contains the results of this pilot study, and the data shows a tendency for the classes taught by the inquiry approach to perform better than those classes taught in a traditional fashion. On the basis of this administration the items were revised. The coefficients of reliability, which were estimated using pretest data with the aid of the Kuder-Richardson Formula 20, are given in Table II.

V. STATISTICAL DESIGN

The fact that random assignment of classes was limited precluded the use of a truly experimental design and allowed, at best, a quasi-experimental one. Campbell and Stanley in discussing a variety of experimental and quasi-experimental designs for research on teaching had this to say about the "Nonequal Control Group Design":

> One of the most widespread experimental designs in educational research involved an experimental group and a control group both given a pretest and a posttest, but in which the control group and the experimental group do not have pre-experimental equivalence. Rather, the groups

TABLE III

PILOT STUDY DATA FOR VALIDATING AND ANALYZING PROCESS MEASURES

GROUP	N	Mean	Standard Deviation
Grade IX - classes taught in a process-oriented fashion	23	62.7	7.5
Grade IX - classes taught in a traditional fashion	32	51.9	6.8
Grade VIII - classes taught in a process-oriented fashion	53	43.5	8.2
Grade VIII - classes taught in traditional fashion	31	39.3	9.1

> constitute naturally assembled collectives such as classrooms, as similar as availability permits but yet not so similar that one can dispense with the pretest (10).

Analysis of Covariance

The techniques of variance and covariance analysis have played major roles in educational research. Since not all the individual units could be assigned at random to the experimental conditions, a covariance analysis was chosen because it incorporates a design which "uses an indirect, or statistical control (1) to increase the precision of the experiment and (2) to remove potential sources of bias in the experiment" (65). Statistical control is achieved by measuring one or more concomitant variables, called covariates, in addition to the variable of primary interest, the criterion. The purpose of the measurement on the covariate is for adjusting the measurements on the criterion.

Although there are many ways in which criterion measures may be adjusted for the influence of the covariates, the most common technique of adjustment is some form of regression analysis. This technique determines the "average effect of an increase of 1 unit in the covariate upon the variate (criterion)" (68). Even though non-linear forms of regression analysis have been used in some instances, the more widely used linear case was used for this study. "In terms of the linear regression, an adjusted criterion mean has the following form,

$$\overline{y'}_{.j} = \overline{y}_{.j} - b(\overline{X}_{.j} - \overline{X}),$$

where b is an estimate of β, the population linear-regression coefficient" (69).

In order to see how differences between adjusted criterion means are freed from the linear effects of the covariate, the basic linear model underlying the analysis of variance has to be examined. Since,

$$\overline{Y}_j = \mu + \beta(\overline{X}_j - \overline{X}) + \tau j + \tilde{\varepsilon} j \ (70),$$

$$\overline{Y}'_j = \mu + \tau j + \tilde{\varepsilon} j,$$

showing that no linear regression effect is present in the difference between two adjusted criterion means.

The advantages of using a one-way analysis of covariance design over a simple investigation of difference scores were thoroughly discussed by Feldt (23).

The design which was used throughout the study is summarized below:

Treatments

Pupil	A			B C D
	Covariates		Criterion	
	COOP test score	pretest score	posttest score	Treatments B, C and D had similar configurations to A.

Figure 2. Design of Analyses used in the Study.

In using this covariance analysis design, a number of assumptions need to be satisfied beyond those of the conventional analyses of variance. These additional assumptions are: first, that treatment effects and regression effects are additive, implying that regressions are homogeneous, and second, that residuals are normally and independently distributed with zero means and the same variance, indicating that the proper form of regression equation has been fitted (71). Fortunately, evidence has shown that F tests in the analysis of covariance "are robust with respect to the violation of the two assumptions, normality and homogeneity of residual variance" (72).

An appropriate F test can be formed in this analysis which takes the form:

$$F = \frac{MS'_{treat}}{MS'_{error}}$$

where: the degrees of freedom for the numerator are (k-1)

the degrees of freedom for the denominator are k(n-1)-1

k = number of treatment groups

n = number of experimental units in each treatment

MS'_{treat} = mean square representing the residual variation about the over-all regression line

MS'_{error} = mean square representing the residual variation about the pooled within-class regression line.

At the University of Alberta the IBM 360 computing facility allows use of a simple and convenient program called ANC010 to do the necessary computation. Throughout the analysis in this study this program was used.

CHAPTER IV

RESULTS AND ANALYSES

This chapter is concerned with the discussion of the analyses to which the data were subjected in order to test the hypotheses set out in Chapter I. As indicated previously, the main analyses were carried out using an analysis of covariance design with two covariates - the COOP test scores and the respective pretest scores. It was thought necessary to use the COOP test score, representing a student's level of achievement in general science, as a covariate in order to somewhat control the variability which may have been incurred due to the limited sampling which was possible. As a further indication of general scholastic ability, IQ scores were recorded. Table IV shows both the Cooperative test score means and IQ score means for each treatment group along with the results of the respective analysis of variance across groups. The table shows that at the beginning of the experiment there was a significant difference among groups on the COOP science test ($p = .005$) and for IQ scores ($p = .016$). The inclusion of the COOP test scores as covariates therefore seems warranted. Although no correlations between these two sets of scores were calculated, an examination of the respective means indicates rather high relationship between the means for all groups except group D which performed much better on the IQ test than on the standardized science test. This may be due in part to the variability in N's, from 125 to 144, causing two rather different samples of students

TABLE IV

MEANS, STANDARD DEVIATIONS AND F-TESTS FOR COOP TEST SCORES AND IQ SCORES

GROUP	COOP TEST			IQ SCORES		
	N	$\overline{X}$	S.D.	N	$\overline{X}$	S.D.
A	137	35.8	9.81	161	114.66	9.71
B	105	38.5	8.74	137	117.71	10.58
C	176	35.3	9.47	178	113.78	10.79
D	125	33.6	11.56	144	115.84	11.31

$F = 4.35$

$p = .005$

$F = 3.46$

$p = .016$

A chi square test showed that the hypothesis of homogeneity of variance is not rejected.

from the tested population of students to be chosen, and partly to the lack of perfect correlation of the COOP science test with IQ measures (the correlation of the COOP general science test - Form B with the SCAT - Parts I and II is given as .62* (22), for example). Both measures were not used as covariates because the IQ scores were not available from the beginning of the study but were released after the main covariance analyses were performed. The difference in the N's accrue either from missing IQ data or from students failing to take the COOP test during the experiment.

Three major sets of hypotheses were postulated: 1.0, 1.1, 1.2 and 1.3 dealing with the "process" dimension, 2.0 dealing with the TOUS test and 3.0 concerned with a more traditional measure of achievement. These three will be discussed separately below.

Hypothesis 1.0

This hypothesis postulated that:

> There will be no difference among the treatment groups in the posttest scores on the total process measure, using the pretest scores on the total process measure and the COOP test scores as covariates.

The results of both the analysis of variance and covariance are recorded in Table V for comparison purposes. The F ratio for the analysis of covariance was computed by using the adjusted within mean square as the denominator. The table shows that the groups differed significantly ($p \leqslant .001$) in both analyses with respect to their performance on the total process measure posttest.

---

*This coefficient represents the correlation between people and not groups means.

TABLE V

TOTAL PROCESS MEASURE ANALYSES

ANALYSIS OF VARIANCE OF POSTTEST SCORES

Source	DF	SS	MS	F	P
Total	478	50,973	. .	. .	. .
Group	3	8,403	2,801.0	31.3	$\ll .001$
Within	475	42,570	89.6	. .	. .

ANALYSIS OF COVARIANCE OF POSTTEST SCORES

Source	DF	MS	Adjusted F	p
Group	3	1,258.0	39.2	$\ll .001$
Within	473	32.1	. .	. .

In order to examine the results further, the unadjusted and adjusted means for the total process measure posttest and the means for the covariance controls are shown in Table VI. The table shows significant increases in scores for groups A, B and C while group D does not seem to change by any significant amount. Group C shows a higher change in means than Groups A and B (+6.2 to +3.7 and +2.9, respectively). That is, the group receiving no particular process emphasis in their instruction attained a higher degree of understanding of processes of science, as measured by the total process measure, than the groups which specifically concentrated on the process of scientific inquiry.

For additional aid in analyzing the results, the intercorrelations among all variables tested were calculated and these are found in Table VII.

In order to test for differences between adjusted treatment means, an average effective error variance per experimental unit needed to be calculated. Winer (73) suggests the following formula:

$$S'^2_{y} = s'^2_{error}\left[1 + \frac{T_{xx}E_{zz} - 2T_{xz}E_{xz} + T_{zz}E_{xx}}{(k-1)(E_{xx}E_{zz} - E^2_{xz})}\right],$$

where S'^2_{error} is the mean square for error in the analysis of covariance (74). The average effective error variance for the difference between two adjusted treatment means then is

$$S'^2_{(\bar{y}'_j - \bar{y}'_k)} = \frac{2S'^2_{y}}{n} \qquad (75).$$

TABLE VI

TOTAL PROCESS MEASURE ANALYSIS

Group	Total Process Posttest Unadjusted	Total Process Pretest	COOP	Total Process Posttest Adjusted
A	41.4	37.7	35.8	41.4
B	43.7	40.8	38.4	40.9
C	44.3	38.1	35.3	43.9
D	33.5	34.1	33.6	36.2
All classes	41.0	37.6	35.6	. .

N	Difference Between Total Process-Post and Total Process-Pre	t	p
125	+3.7	9.40	$<.001$
82	+2.9	7.80	$<.001$
164	+6.2	15.40	$\ll.001$
108	-0.6	0.89	N.S.

TABLE VII

INTERCORRELATIONS AMONG VARIABLES

N = 568

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Coop	1.00	0.59	0.59	0.45	0.54	0.60	0.67	0.51	0.53	0.65	0.70	0.62	0.69
2. Process (I) Pre		1.00	0.49	0.43	0.43	0.63	0.59	0.55	0.46	0.53	0.55	0.76	0.56
3. Process (II) Pre			1.00	0.28	0.32	0.50	0.67	0.45	0.35	0.46	0.54	0.73	0.40
4. Process (III) Pre				1.00	0.40	0.42	0.46	0.34	0.29	0.52	0.36	0.71	0.34
5. Tous-Pre					1.00	0.44	0.44	0.34	0.64	0.46	0.46	0.31	0.36
6. Process (I) Post						1.00	0.57	0.48	0.43	0.46	0.62	0.41	0.83
7. Process (II) Post							1.00	0.50	0.51	0.57	0.68	0.58	0.81
8. Process (III) Post								1.00	0.37	0.46	0.59	0.40	0.76
9. Tous-Post									1.00	0.45	0.52	0.39	0.53
10. Achievement Pre										1.00	0.66	0.57	0.53
11. Achievement Post											1.00	0.56	0.67
12. Total Process (Pre)												1.00	0.68
13. Total Process (Post)													1.00

Unfortunately, the output of ANCO 10 at the University of Alberta did not make the exact calculation of S'^2_y possible. However, in an effort to have some indication of how significant these differences might be, S'^2_y was allowed to vary from S'^2_{error} to 1.5 S'^2_{error}. Even upon calculation it is almost certain to be within these limits. That is, the term $\frac{T_{xx}E_{zz} - 2T_{xz}E_{xz} + T_{zz}E_{zz}}{(k-1)\ (E_{xx}E_{zz} - E^2_{xz})}$ was allowed to vary from 0.0 to 0.5. An examination of a number of sample analyses showed this term never to vary beyond the above limits. To further insure the reasonableness of these limits, a sample analysis (N=20) was done in order to calculate a value for the above term. The value determined was 0.024 and the whole analysis is reproduced in Appendix B. Table VIII shows the calculation of F ratios using both average effective error variances. Clearly, differences between treatment means A and C, A and D, B and C, B and D, and C and D are significant beyond the 0.01 probability levels while the differences between the means of A and B is not significant.

In the above analysis, it can be seen that, after taking into account initial differences in understanding and differences in general science achievement, groups A, B and C, the groups emphasizing processes of science and the group receiving traditional instruction in the same content area, attained a significantly greater understanding of science, as measured by the total process measure, than students who did not. Moreover, after taking the above differences into account, Group C

TABLE VIII

TESTING FOR DIFFERENCES IN ADJUSTED GROUP MEANS ON TOTAL PROCESS MEASURE

$\tilde{n}^* = 113$

	1. Using $S'^2_y = S'^2_{error}$ (32.1)				2. Using $S'^2_y = 1.5S'^2_{error}$ (48.2)			
GROUP	A	B	C	D	A	B	C	D
$\bar{y}'$	41.4	40.9	43.9	36.2	41.4	40.9	43.9	36.2
	$F^{***}_{A,B}$	=	.4		F	=	.3	
	$F_{A,C}$	=	11.0	**	F	=	7.3	**
	$F_{A,D}$	=	42.5	**	F	=	31.7	**
	$F_{B,C}$	=	15.8	**	F	=	10.6	**
	$F_{B,D}$	=	38.9	**	F	=	25.8	**
	$F_{C,D}$	=	104.0	**	F	=	69.6	**

$F_{.95(1,\infty)} = 3.84$

$F_{.99(1,\infty)} = 6.63$

---

*Since n's were not equal, a harmonic mean $\tilde{n}$ was calculated as suggested by Winer (67).

**Indicates significance at .01 level of significance.

***Where $F_{A,B} = \dfrac{(\bar{y}'_A - \bar{y}'_B)^2}{\dfrac{2\,S'^2_y}{\tilde{n}}}$

performed significantly better than all other treatment groups.

In view of the above findings, it may be well to describe the treatment groups again. Treatment A emphasized the processes of science and the teachers participated in the preparation of the curricular materials. Treatment B also emphasized the process dimension but the teachers were handed prepared materials and in no way participated in their preparation. Treatment group C taught the same content in a "traditional" fashion while treatment D served as a control group receiving no instruction in the experimental content area. As was indicated earlier, however, students and teachers in treatment C did encounter the Inventory, but it was not expected that they would use it in their learning and teaching, respectively.

To further explore the relative effectiveness of the various methods for the teaching of the process dimension of science, the total process measure was broken down into its three constituent parts (I, II and III), each dealing with a different aspect of the process dimension. Part I dealt with knowledge and understanding of the processes in scientific inquiry as shown by tests based on investigations performed in class. Part II was an attempt to measure a student's understanding and ability to apply the processes of science in similar content areas but not previously encountered experiments. Part III was an attempt to determine how well a student was able to generalize his knowledge and understanding of the processes of scientific inquiry by confronting him with physical science experiments not directly related to the experimental content. The analyses based on each of these three parts follow below and

are based on hypotheses 1.1, 1.2 and 1.3 respectively. Hopefully, a more detailed examination of the process dimension may shed more light on the above results.

Hypothesis 1.1

This hypothesis postulated that:

> There will be no difference among the treatment groups in the posttest scores on the process measure (Part I), using the pretest scores on the process measure (Part I) and the COOP test scores as covariates.

The results of both the analysis of variance and covariance for Part I (Comprehension) of the process measure are recorded in Table IX. The F ratio for the analysis of covariance was computed by using the adjusted within mean square as the denominator. The table shows that the groups differed significantly ($p \ll .001$) in both analyses with respect to their performance on Part I of the process measure posttest.

The unadjusted and adjusted means for Part I of the process measure posttest and the means for the covariance controls are shown in Table X. The significant difference between means on Part I of the process measure posttest is clearly in favor of the three experimental treatments while control group D showed no gain. The data in the lower half of Table X shows gains in Part I of the process measure scores from pretest to unadjusted posttest. The respective gains clearly are in favor of group C, group A gaining the least (+1.7) among the experimental groups.

Table XI shows the tests for the difference among adjusted posttest means using the same rationale for the calculation of the effective error variance as in the total process measure analysis (p.64).

TABLE IX

PROCESS MEASURE (PART I) ANALYSES

Analysis of Variance of Posttest Scores

Source	DF	SS	MS	F	P
Total	532	10,184	. .	. .	. .
Group	3	1,785	595.0	37.5	$\ll .001$
Within	529	8,399	15.9	. .	. .

Analysis of Covariance of Posttest Scores

Source	DF	MS	F	P
Group	3	377.0	42.7	$\ll .001$
Within	527	8.8	. .	. .

TABLE X

PROCESS MEASURE (PART I) ANALYSIS

Group	Posttest Unadjusted	Pretest	Coop	Posttest Adjusted
A	14.1	12.4	35.8	14.2
B	16.2	14.1	38.4	15.0
C	16.1	12.8	35.3	16.1
D	11.5	11.9	33.6	12.1
All classes	14.6	12.7	35.6	. .

N	Difference between Post and Pretest	t	p
137	+1.7	4.91	.001
105	+2.1	5.67	$<.001$
176	+3.3	8.50	$<.001$
115	-0.4	0.15	N.S.

Clearly, differences between treatment means A and C, A and D, B and D, and C and D are significant beyond the 0.01 probability level even when tested using the rather conservative 1.5 S'^{2}_{error} for the calculation of the effective error variance. Differences between adjusted means of treatment groups A and B are significant somewhere between the 0.05 and 0.01 probability levels, group B outperforming group A slightly. Group B is significantly different from Group C using S'^{2}_{error} as an estimate for the effective error variance.

In the above analysis it can be seen that, after taking into account initial differences in understanding and differences in general science achievement, groups A, B and C attained a greater understanding of science, as measured by Part I of the process measure, than group D. More specifically, after taking into account intial differences, group C, those classes receiving the so-called traditional instruction, attained a significantly higher degree of understanding of the scientific enterprise, as measured by Part I of the process measure, than those groups specifically instructed in the process of science, A and B - a somewhat unexpected result.

<u>Hypothesis 1.2</u>

This hypothesis postulated that:

> There will be no difference among the treatment groups in the posttest scores on the process measure (Part II), using the pretest scores on the process measure (Part II) and the COOP test scores as covariates.

The results of both the analysis of variance and covariance for

TABLE XI

TESTING FOR DIFFERENCES IN ADJUSTED GROUP MEANS
PROCESS MEASURE (PART I)

$\tilde{n}$* = 113

	Using $S'^2_y = S'^2_{error}$ (8.8)				Using $S'^2_y = 1.5\ S'^2_{error}$ (13.2)			
GROUP	A	B	C	D	A	B	C	D
$\bar{y}'$	14.2	15.0	16.1	12.1	14.2	15.0	16.1	12.1
$F_{A,B}$		5.0				3.3		
$F_{A,C}$		28.0	**			18.7	**	
$F_{A,D}$		34.0	**			22.8	**	
$F_{B,C}$		9.4	**			6.3		
$F_{B,D}$		65.0	**			43.5	**	
$F_{C,D}$		124.0	**			83.0	**	

$F_{.95}(1,\infty) = 3.84$

$F_{.99}(1,\infty) = 6.63$

*Since n's were not equal, a harmonic mean, $\tilde{n}$, was calculated as suggested by Winer (67).

**Indicates significance at .01 level of significance.

Part II (application) of the process measure are recorded in Table XII. The F ratio for the analysis of covariance was computed by using the adjusted within mean square as the denominator. The table shows that the groups differed significantly ($p < .001$) in both analyses with respect to their performance on Part II of the process measure posttest.

The unadjusted and adjusted means for Part II of the process measure posttest and the means for the covariance controls are shown on the top of Table XIII. The significant difference between means on Part II of the process measure posttest is clearly in favour of the three experimental groups, all three having roughly identical unadjusted posttest means.

The difference scores at the bottom of the table, however, show significant gains in Part II of the process measure scores for groups A, C and D while the slight gain by group B is not significant. A quick glance at the pretest scores shows that group B may have been at a slight disadvantage by scoring rather high on the pretest.

Table XIV shows the tests for the differences among adjusted posttest means using a similar rationale for the calculation of the effective error variance to that used in the total process measure analysis. Differences between adjusted treatment means A and D, B and C, and C and D are clearly significant beyond the 0.01 probability level even when 1.5 S'^{2}_{error} is used as an estimate for the effective error variance. All other differences among adjusted means are not significant at the 0.05 probability level.

TABLE XII

PROCESS MEASURE (PART II) ANALYSES

Analysis of Variance of Posttest Scores

Source	DF	SS	MS	F	P
Total	534	9,415	. .	. .	. .
Group	3	768	256.0	15.7	$<.001$
Within	531	8,647	16.3	. .	. .

Analysis of Covariance of Posttest Scores

Source	DF	MS	F	P
Group	3	90.2	11.3	$<.001$
Within	529	8.0	. .	. .

TABLE XIII

PROCESS MEASURE (PART II) ANALYSIS

Group	Posttest Unadjusted	Pretest	COOP	Posttest Adjusted
A	12.7	11.1	35.8	12.6
B	12.3	12.0	38.4	12.0
C	12.6	10.8	35.3	13.1
D	10.0	9.0	33.6	11.3
All classes	11.4	11.0	35.6	. .

N	Difference between Post and Pretest	t	p
137	+1.6	4.39	$<.001$
107	+0.3	.59	N.S.
175	+1.8	6.48	$<.001$
117	+1.0	3.21	.002

TABLE XIV

TESTING FOR DIFFERENCES IN ADJUSTED GROUP MEANS
PROCESS MEASURE (PART II)

$\tilde{n}$* = 113

	Using $S'^2_y = S'^2_{error}$ (8.0)				Using $S'^2_y = 1.5\ S'^2_{error}$ (12.0)			
GROUP	A	B	C	D	A	B	C	D
$\bar{y}'$	12.6	12.0	13.1	11.3	12.6	12.0	13.1	11.3
$F_{A,B}$		3.0				2.0		
$F_{A,C}$		2.1				1.4		
$F_{A,D}$		14.2	**			9.5	**	
$F_{B,C}$		10.2	**			6.8	**	
$F_{B,D}$		4.1				2.7		
$F_{C,D}$		27.1	**			18.1	**	

$F_{.95(1,\infty)} = 3.84$

$F_{.99(1,\infty)} = 6.63$

*Since n's were not equal, a harmonic mean, $\tilde{n}$, was calculated as suggested by Winer (67).

**Indicates significance at .01 level of significance.

In the above analysis it can be seen that, after taking into account intial differences in understanding and differences in general science achievement, the experimental groups A, B and C attained a greater understanding of science, as measured by Part II of the process measure, than group D. More specifically, after taking into account initial differences, the order of achievement was group C, A and B, respectively, the respective intervals being roughly equal. That is, group C, those classes receiving the so-called "traditional" instruction, attained a higher (only $F_{B,C}$ being significant) understanding of the scientific enterprise, as measured by Part II of the process measure, than those groups specifically instructed in the process of science- A and B.

Hypothesis 1.3

This hypothesis postulated that:

> There will be no difference among the treatment groups in the posttest scores on the process measure (Part III), using the pretest scores on the process measure (Part III) and the COOP test scores as covariates.

The results of both the analysis of variance and covariance for Part III (transfer) of the process measure are recorded in Table XV. The F ratio for the analysis of covariance was computed by using the adjusted within mean square as the denominator. The table shows that the groups differed significantly ($p < .001$) in both analyses with respect to their performance on Part III of the process measure posttest.

TABLE XV

PROCESS MEASURE (PART III) ANALYSES

Analysis of Variance of Posttest Scores

Source	DF	SS	MS	F	p
Total	507	6,590	. .	. .	. .
Group	3	701	234.0	20.0	$<.001$
Within	504	5,889	11.7	. .	. .

Analysis of Covariance of Posttest Scores

Source	DF	MS	F	p
Group	3	136.0	16.9	$<.001$
Within	502	8.1	. .	. .

The unadjusted and adjusted means for Part III of the process measure posttest and the means for the covariance controls are shown on the top of Table XVI. The significant difference between means on Part III of the process measure posttest is clearly in favour of the three experimental groups, all three having roughly identical unadjusted posttest means.

The difference scores at the bottom of Table XVI more clearly show this result along with a significant negative difference in means for group D.

Table XVII shows the tests for the differences among adjusted posttest means using a similar rationale for the calculation of the effective error variance to that used in the total process measure analysis. Differences between adjusted treatment means A and D, B and D, and C and D are clearly significant beyond the 0.01 probability level even when 1.5 S'^2_{error} is used as an estimate for the effective error variance. All other differences among adjusted means are not significant at the 0.01 probability level.

In the above analysis it can be seen that, after taking into account initial differences in understanding and differences in general science achievement, the experimental groups, A, B and C attained a greater understanding of science, as measured by Part III of the process measure, than group D.

As a matter of fact, the differences among adjusted treatment means of groups A, B and C were negligible. That is, all three

TABLE XVI

PROCESS MEASURE (PART III) ANALYSIS

Group	Posttest Unadjusted	Pretest	COOP	Posttest Adjusted
A	14.6	13.9	35.9	14.6
B	14.7	14.5	38.5	14.3
C	14.8	14.3	35.1	14.5
D	11.9	13.0	33.2	12.5
All classes	14.0	13.9	35.6	. .

N	Difference between Post and Pretest	t	p .
136	+0.7	1.31	N.S.
90	+0.2	.52	N.S.
167	+0.5	.79	N.S.
115	-1.1	2.74	.003

TABLE XVII

TESTING FOR DIFFERENCES IN ADJUSTED GROUP MEANS
PROCESS MEASURE (PART III)

$\tilde{n}$* = 113

	Using $S'^2_y = S'^2_{error}$ (8.1)				Using $S'^2_y = 1.5\ S'^2_{error}$ (12.2)			
GROUP	A	B	C	D	A	B	C	D
$\bar{y}'$	14.6	14.3	14.5	12.5	14.6	14.3	14.5	12.5
$F_{A,B}$		.7				.4		
$F_{A,C}$		.08				.05		
$F_{A,D}$		36.2	**			24.1	**	
$F_{B,C}$		.3				.2		
$F_{B,D}$		26.6	**			17.7	**	
$F_{C,D}$		33.0	**			21.9	**	

$F_{.95}(1,\infty) = 3.84$

$F_{.99}(1,\infty) = 6.63$

*Since n's were not equal, a harmonic mean, $\tilde{n}$, was calculated as suggested by Winer (67).

**Indicates significance at .01 level of significance.

experimental groups receiving some instruction in the "Matter and Energy" content area performed equally well on Part III of the process measure after initial differences had been taken into account.

Hypothesis 2.0

This hypothesis postulated that:

> There will be no difference among the treatment groups in the posttest scores on the TOUS, using the TOUS pretest scores and the COOP test scores as covariates.

The results of both the analysis of variance and covariance for the TOUS are recorded in Table XVIII. The F ratio for the analysis of covariance was computed by using the adjusted within mean square as the denominator. The table shows that the groups differed significantly ($p < .001$) in both analyses with respect to their performance on the TOUS posttest.

Table XIX displays the unadjusted and adjusted means for the TOUS posttest and the means for the covariance controls. The significant difference between unadjusted TOUS posttest means is clearly in favour of the three experimental treatments while group D remained relatively stable. The lower half of Table XIX provides a more detailed analysis of the difference between pre and posttest means. Although groups A and C showed some gain, only group B showed a significant difference between pre and posttest means.

The tests for the differences among adjusted posttest means are shown in Table XX. The rationale for the calculation of the effective error variance was the same as that use in the total process measure

TABLE XVIII

TOUS ANALYSES

Analysis of Variance of Posttest Scores

Source	DF	SS	MS	F	P
Total	566	13,785	. .	. .	. .
Group	3	1,785	595.0	27.9	$\ll .001$
Within	563	12,000	21.3	. .	. .

Analysis of Covariance of Posttest Scores

Source	DF	MS	F	p
Group	3	158.8	15.4	$< .001$
Within	561	10.3	. .	. .

TABLE XIX

TOUS ANALYSIS

Group	Posttest Unadjusted	Pretest	COOP	Posttest Adjusted
A	21.2	20.5	35.9	21.7
B	23.5	21.8	38.5	23.4
C	21.5	20.6	35.1	22.6
D	18.8	19.2	33.2	20.7
All classes	22.1	21.3	35.6	. .

N	Difference between Post and Pretest	t	p
151	+0.7	1.57	N.S.
112	+1.7	2.83	.005
182	+0.9	1.72	N.S.
119	-0.4	0.75	N.S.

TABLE XX

TESTING FOR DIFFERENCES IN ADJUSTED GROUP MEANS
TOUS

$\tilde{n}^* = 140$

	Using $S'^2_y = S'^2_{error}$ (10.3)				Using $S'^2_y = 1.5\ S'^2_{error}$ (16.0)			
GROUP	A	B	C	D	A	B	C	D
$\bar{y}'$	21.7	23.4	22.6	20.7	21.7	23.4	22.6	20.7
$F_{A,B}$		13.3	**			8.6	**	
$F_{A,C}$		5.5				3.5		
$F_{A,D}$		6.8	**			4.5		
$F_{B,C}$		4.4				2.8		
$F_{B,D}$		49.6	**			31.9	**	
$F_{C,D}$		24.5	**			15.8	**	

$F_{.95(1,\infty)} = 3.84$

$F_{.99(1,\infty)} = 6.63$

*Since n's were not equal, a harmonic mean, $\tilde{n}$, was calculated as suggested by Winer (67).

**Indicates significance at .01 level of significance.

analysis.

The double asterisks indicate significant differences at the 0.01 probability levels between treatment means A and B, B and D, and C and D while $F_{A,D}$ is probably significant somewhere between the 0.01 and 0.05 probability levels. Of special significance or interest is the significant difference between treatment groups A and B.

In the above analysis it can be seen that, after taking into account initial differences in understanding and differences in general science achievement, groups A, B and C attained a greater understanding of science, as measured by the TOUS, than group D. More specifically, after taking initial differences into account, group B performed best among the experimental groups (groups C, A and D following in that order).

Hypothesis 3.0

This hypothesis postulated that:

> There will be no difference among the treatment groups with respect to the posttest scores on the achievement test, using the achievement pretest scores and the COOP test scores as covariates.

The results of both the analysis of variance and covariance for the achievement test scores are recorded in Table XXI. The F ratio for the analysis of covariance was computed by using the adjusted within mean square as the denominator. The table shows that the groups differed significantly ($p < .001$) in both analyses with respect to their performance on the achievement test.

TABLE XXI

ACHIEVEMENT TEST ANALYSES

Analysis of Variance of Posttest Scores

Source	DF	SS	MS	F	p
Total	587	22,143	. .	. .	. .
Group	3	2,414	804.6	23.8	< .001
Within	584	19,729	33.8	. .	. .

Analysis of Covariance of Posttest Scores

Source	DF	MS	F	p
Group	3	469.0	31.4	< .001
Within	582	14.9	. .	. .

The unadjusted and adjusted means for the achievement posttest and the means for the covariance controls are shown in Table XXII. The significant difference between means on the posttest is clearly in favour of the three experimental treatments while control group D showed no gain. The bottom of the table shows more specifically that all three experimental treatment groups (A,B,C) made significant gains in achievement scores from pretest to unadjusted posttest.

Table XXIII shows the tests for the differences among adjusted posttest means using the same rationale for the calculation of the effective error variance as in the total process measure analysis. Clearly, differences between treatment means A and D, B and D, and C and D are significant beyond the 0.01 probability level even when tested using the rather conservative 1.5 S'^{2}_{error} for the calculation of the effective error variance. Differences between adjusted means of treatment groups B and C are significant somewhere between the 0.05 and 0.01 probability levels, group C outperforming group B slightly.

In the above analysis it can be seen that, after taking into account initial differences in understanding and differences in general science achievement, groups A, B and C attained a greater understanding of the science content studied, as measured by the constructed achievement test, than group D. More specifically, after taking into account initial differences, group C, those classes receiving the so-called "traditional" instruction, attained a higher degree of understanding of the "Matter and Energy" content, as measured by the achievement test, than groups A and B which were also instructed in the same content area, though the coverage may not necessarily have been the same.

TABLE XXII

ACHIEVEMENT TEST ANALYSIS

Group	Posttest Unadjusted	Pretest	COOP	Posttest Adjusted
A	20.3	16.7	35.9	20.3
B	21.6	18.8	38.5	20.0
C	21.1	16.8	35.1	21.2
D	16.2	16.6	33.2	17.0
All classes	19.8	17.1	35.6	. .

N	Difference between Post and Pretest	t	p
151	+3.6	10.69	$<.001$
112	+2.8	7.08	$<.001$
182	+4.3	12.30	$<.001$
119	-0.4	1.00	N.S.

TABLE XXIII

TESTING FOR DIFFERENCES IN ADJUSTED GROUP MEANS ON ACHIEVEMENT TEST

$\tilde{n}$* = 150

	Using $S'^2_y = S'^2_{error}$ (14.9)				Using $S'^2_y = 1.5\ S'^2_{error}$ (22.5)			
GROUP	A	B	C	D	A	B	C	D
$\bar{y}'$	20.3	20.0	21.2	17.0	20.3	20.0	21.2	17.0
$F_{A,B}$		.5				.3		
$F_{A,C}$		4.1				2.7		
$F_{A,D}$		55.0	**			36.7	**	
$F_{B,C}$		7.2	**			4.8		
$F_{B,D}$		45.0	**			30.0	**	
$F_{C,D}$		88.0	**			58.8	**	

$F_{.95(1,\infty)} = 3.84$

$F_{.99(1,\infty)} = 6.63$

*Since n's were not equal, a harmonic mean, $\tilde{n}$, was calculated as suggested by Winer (67).

**Indicates significance at .01 level of significance.

CHAPTER V

SUMMARY, CONCLUSION, DISCUSSIONS, AND RECOMMENDATIONS

I. SUMMARY

In the past few years a strong movement has become noticeable to change the emphasis in the teaching of science. This change in emphasis has to do with the objectives of science teaching and has been the result, largely, of careful study and investigation of the definition and nature of science -especially the work of Schwab and Brandwein. Science is no longer regarded merely as a "body of systematized knowledge," but also as a "process of inquiry, resulting in a body of systematized knowledge."

As a result, the two major objectives of science education are for the child to develop (1) knowledge of science concepts - the content or product of science - and (2) facility in scientific skills - the processes of scientific inquiry. For many years objectives stated by various agencies have urged this dual emphasis in science teaching. However, all too often the major emphasis has been on content, and process objectives have been ignored. This has effected expositional teaching and "memorizational learning" on the part of the students, with an accompanying deprivation of needed experiences with scientific inquiry and the processes of science.

Understandably, then, there is presently a widespread preoccupation with developing new programs emphasizing the process approach to the learning of science.

This study attempted the development of a science curriculum based on both the substance and syntax of science, and is only the first in a series of longitudinal studies directed at investigating science teaching from this point of view. Specifically, curriculum materials were developed for a unit of science content, at the grade VII or VIII level, which recognized the importance of the process dimension in science teaching and learning. As such, a conscientious effort was made to incorporate the processes of scientific inquiry into student investigations wherever possible. A helpful contribution in this regard was the development of a theoretical framework, An Inventory of Processes in Scientific Inquiry, which (1) guided the programming of the proposed unit, (2) was indicative of special teaching strategies, and (3) served as a framework for evaluation.

A number of grade VII and VIII science classes were assigned to one of four different treatment groups - the treatments differing in the emphases that were placed on the integration of the process dimension in the teaching and the involvement of the respective teachers in the development of the curriculum materials. That is, the treatments ranged from (1) incorporating special strategies for the teaching for the processes of scientific inquiry into materials which teachers about to teach these materials helped to prepare (Treatment A),(2) incorporating the same strategies for teaching for the process objectives but with no participation on the part of teachers (Treatment B), (3) teaching the same content in a so-called traditional fashion (Treatment C), and (4)

teaching no related content unit in any special manner (Treatment D).

Control variables, used to aid in the interpretation of the analyses, included: a general science ability and achievement test (the co-operative general science test, Form B); a test of achievement related to the content of investigations performed (the Achievement Test); a test measuring understanding of the scientific enterprise, the processes of scientific inquiry and characteristics of scientists (the TOUS); and tests measuring understanding of the processes of scientific inquiry at various levels (the process measures). All tests, except the COOP test, were administered in random order before and after the experimental period.

II. CONCLUSIONS AND DISCUSSION

The order of the conclusions and their respective discussions is the same as that of the hypotheses stated in Chapter I. Overall however, the conclusions are divided into two broad categories, those directly related to the hypotheses, and those related to the general conclusions.

A. Conclusions Related to Performance on Tests.

Total Process Measure. Hypothesis 1.0 postulated no difference among treatment groups with respect to their post-test scores on the total process measure, using the pretest scores on the total process measure and the COOP test scores as controls. The results as indicated on Tables V,VI, and VIII lead to a rejection of the null hypothesis. Students in the various treatment groups did differ in their performance on the total process measure post-test in the presence of the above-

mentioned controls.

Clearly, treatment group C performed significantly better than all other treatment groups and showed a surprisingly high difference score between pretest and unadjusted posttest scores. Treatment C's performances on the three subtests (Parts I, II and III) of the total process measure revealed gain scores of +3.3, +1.8 and +0.5, respectively.* The correlation** between Parts I, II and III of the process measure posttests and the total process measure posttest were 0.83, 0.81 and 0.76, respectively.

The respective gain scores are quite plausible when interpreted in the light of the hierarchy of meaningfullness which was postulated earlier and which led to the inclusion of three process tests to measure understanding of the process dimension at three levels. Part I was a measure of how well students understood the processes in scientific inquiry in investigations they performed in class. It seems that, even though treatment group C only encountered the designated investigations by demonstration accompanied by, perhaps, verbal explanations, they performed significantly better on Part I of the process measure than students who actually performed the said investigations individually or in pairs. It should be noted here also that the classes in Treatment C covered more of the prepared curriculum than did Treatments A and B.

---

*See Tables X, XIII and XVI.

**See Table VII

The additional content covered may have significantly affected their scores on the process measures.

In addition, this seemingly anomalous performance by treatment C may be partially accounted for by the verbal nature of the testing instrument (even though visual stimuli were used). Perhaps the manner in which investigations were demonstrated or explained in the traditional treatment group may have actually caused some of the test questions and their answers to be verbalized. This may or may not have been the case in treatment groups A and B because of the individual or small group nature of the learning situation. This also brings up the interesting possibility that as long as these rather verbal measures are used for testing process objectives, pupils instructed verbally in the processes (as opposed to students <u>using</u> processes in the laboratory situation) perform better on these measures. That is, perhaps the constructed process measures were not as valid as initially assumed even though they seemed to display adequate face validity. They may not be able to discriminate among treatment groups receiving varied instruction in the processes of scientific inquiry. They may, in fact, favor verbal instruction.

An additional consideration dealing with the validity of the process measures is that while these measures appeared to have face validity (because they were based on the Inventory), they may not necessarily be valid measures in light of the instruction that students encountered. While it is relatively easy to incorporate processes of scientific inquiry into a content area, it is much more difficult to develop strategies of <u>teaching</u> by the process approach. There may be an overlap here but the two phenomena are not synonymous. The point is that the various teachers

formulated their <u>own</u> approach to the teaching of the processes incorporated into the prepared materials. There are indications (discussions and anecdotal records) that no two approached this problem in the same manner. In addition, the writer made no attempt to be secretive about the nature of this study and participating teachers in group C, especially, may have unintentionally emphasized particular processes in their teaching from time to time.

An interesting sidelight was the rather high pretest score attained by treatment group B on the total process measure, putting them at a slight disadvantage from the start by making it more difficult for them to show a large gain in unadjusted scores. This phenomenon may have been a possible effect of the large number of "new science curricula" these students had recently encountered and also because all these pupils were in grade VIII.

Another observation, which will be discussed in more detail in the section dealing with achievement test results, deals with the almost parallel performance of the various treatment groups on the process measures and the achievement test.* Suffice it to say here that both measures were purposely constructed in such a way that predominantly higher mental process objectives were measured.**

Quite independently from some of the above discussion and conclusions, it should be noted that the total process measure was indeed able to discriminate between treatments which used the "Matter and Energy" content unit and those that did not - even though it did not provide outcomes in the anticipated direction as far as discrimination between the experimental and traditional treatments were concerned.

---

*See Tables VI and XXII.

**See discussion on pp. 10-11.

A number of reasons explaining the obtained results have been discussed; one of the most crucial, though, as far as the process measure is concerned, has not been mentioned. The reference is, of course, to the short, two-month duration of the experiment. It seemed as if the students and teachers in treatment groups A and B barely had a chance to reorient themselves to this novel approach to science learning and teaching when the experiment was over. Treatment C, on the other hand, presumably continued in a manner quite similar to that encountered by students and teachers previous to the experiment.

A related point deals with the amount that the teaching in the various treatment groups was actually controlled. It may be that the teachers in group A felt very restricted in their teaching approach while groups B and C were given complete freedom in their approach to the curriculum materials. Another way of stating this is that the Hawthorne Effect may have been greater for the teachers in treatments B and C than those in group A.

Hopefully, a more detailed discussion of the three subtests of the total process measure will shed more light on some of the above conclusions.

Process Measure - Part I. Hypothesis 1.1 postulated no difference among treatment groups with respect to their posttest scores on Part I of the process measure, using the pretest scores on the process measure (Part I) and the COOP test scores as controls. The results displayed on Tables IX, X and XI lead to a rejection of the null hypothesis. Students in the four treatment groups did not perform equally well on Part I of the

process measure - an instrument designed to measure comprehension of processes of scientific inquiry in investigations performed as part of the experimental and traditional curriculum.

As indicated previously, Treatments A, B and C showed highly significant gains from pretest to unadjusted posttest scores while Treatment D remained relatively stable. In terms of adjusted posttest scores, differences in favour of Treatment C are not as pronounced here as they were for the total process measure.

It seems that if tests are based directly on proceedings which occurred in the classroom, even if these tests are of the modified paper-and-pencil variety, then students using the inquiry or process-oriented approach as defined by the experimental curriculum have no advantage (and perhaps a slight disadvantage) on these tests. That is, whether a certain test item, even on the process measure, measures recall or a higher mental process, depends almost entirely on whether the student has encountered that specific process <u>and</u> content in the learning situation. Hence the situation here is very similar to classifying items according to Bloom's taxonomy in more traditional tests. The above discussion points to the futility and error of classifying items, in terms of Bloom's taxonomy, before the actual teaching occurs. Perhaps this procedure of item categorization can only be done a priori if it is certain that the student will not encounter the content used in the item. As pointed out earlier the relatively high achievement by treatment C may be due to the fact that these students encountered more of the experimental curriculum than did groups A and B. It is possible that content is more closely linked to the process dimension

than is usually admitted, and that it is unlikely that the two dimensions can be completely separated for testing purposes. Analyzing the results of the posttests of Part II and III of the process measure may provide more evidence in this regard.

Process Measure - Part II. Hypothesis 1.2 postulated no difference among treatment groups with respect to their posttest scores on the process measure (Part II), using the pretest scores on Part II of the process measure and the COOP test scores as controls. On the basis of the analyses performed and the results found in Tables XII, XIII and XIV, the null hypothesis is rejected. It must be emphasized, however, that this rejection is due mainly to the fact that group D differed significantly from the others - a somewhat expected and anticipated result. All treatment groups showed increases from pretest to unadjusted posttest; the increase for group B (+0.3) was not, however, significant.* An increase which is difficult to interpret is that displayed by group D (+1.0). In only this one sub-test did a significant increase occur for this group.

In light of the discussion of Part I immediately preceding this section it is interesting to note that the adjusted posttest scores on Part II of the process measure, an instrument designed to measure a student's ability to comprehend and use the processes of scientific inquiry with content related to the content covered in class but not identical to it, do not exhibit the same marked trends in favour of treatment group C. Only the difference between treatments B and C was

*See Table XIII.

significant at the 0.05 level of significance *, aside from the significant AD and CD differences. Attention is drawn once again to the high pretest score obtained by treatment B.

Regarding the argument of "content" confounding the measurement of process objectives, there appears to be evidence here that when "content" objectives are removed from tests, differences in performance on "process" measures among groups are not as pronounced.

Process Measure - Part III. Hypothesis 1.3 postulated no difference among treatment groups with respect to their posttest scores on the process measure (Part III), using the pretest scores on the process measure (Part III) and the COOP test scores as covariates. On the basis of results tabulated on Tables XV, XVI and XVII, hypothesis 1.3 is rejected. Clearly, there were insignificant gains between pretest and unadjusted posttest scores for treatment groups A, B and C while group D showed a significant loss.**

The difference among adjusted posttest scores is negligible in all cases except between treatment group D and all other groups. This part of the process measure was designed to measure a student's ability to apply the processes in scientific inquiry in the area of general science only incidentally related to the area under study. The fact that the test did not discriminate between treatments emphasizing the process approach (A and B) and the treatment using the more traditional approach (C) probably is a direct result of the short encounter (2 months)

---

*See Table XIV.

**See Table XVI, p. 79.

which treatment group A and B had with the process approach before posttests were administered. This severe time limitation applies euqally well to the preceding discussions related to hypothesis 1.0, 1.1 and 1.2, and is probably the main reason for the disappointing failure of the process measure as a whole to discriminate between treatments A and B on the one hand and treatment C on the other.

TOUS (Test on Understanding Science). Hypothesis 2.0 postulated no difference among treatment groups in the posttest scores on the TOUS, using the TOUS pretest scores and the COOP test scores as covariates. The results lead to a rejection of the null hypothesis*. In terms of difference scores between pretest and unadjusted posttest, only group B showed a significant increase, all other changes being negligible. The adjusted posttest scores reveal a surprisingly high mean for treatment group B causing a significant difference between group A and B** in adjusted posttest scores. This may be due in part to the fact that all classes in treatment B were grade VIII classes and perhaps to the fact that this treatment group had the highest mean IQ (117.71)***. Hukins (29), in a factorial analysis of various science tests claimed that the TOUS has high loadings on a verbal factor, this factor usually being highly correlated with most verbal IQ measures used in the schools.

It is significant to point out that Treatments A, B and C did show positive gains while group D had a slight loss. This may indicate

---

*See Tables XVIII, XIX, and XX.

**See Table XX.

***See Table IV, p. 58.

some relation between the unit under study and the TOUS, even though correlations with other tests tended to be generally low*. The lack of high correlation between the TOUS and the process measures can be partially attributed to results obtained from an examination of the items in the TOUS. A majority of items deal with science as a human enterprise, communication among scientists, science and society, scientists as people, institutional pressures on scientists, etc., while a definite minority of items deal with science as a process of inquiry - the focus of this study.

Achievement Test. Hypothesis 3.0 postulated no difference among treatment groups with respect to the posttest scores on the achievement test, using the achievement pretest scores and the COOP test scores as covariates. The results presented in Tables XXI, XXII and XXIII lead to a rejection of the null hypothesis. Treatments A, B and C showed clearcut gains from pretest to unadjusted posttest; group D once again remaining relatively stable. The analysis of differences between adjusted posttest means** clearly illustrates that the three treatments of import (A, B and C) did not differ significantly. This result lends support to a number of similar results in studies showing gains in content-oriented objectives when inquiry or process-objectives were emphasized in the teaching (34). Examples of these types of studies are those comparing laboratory versus non-laboratory instruction, or inductive versus "cookbook" laboratory instruction. That is, though processes in scientific inquiry

---

*See Table VII.

**See Table XXIII.

were emphasized in the teaching of treatment groups A and B, this emphasis resulted in only slightly poorer performance on the achievement test for these groups. This result is especially significant when it is considered that the classes in treatment C covered more content than group A or B. It should also be noted that the curriculum materials prepared for treatment C contained more content background than did those of treatments A and B. In light of the above arguments, the slightly superior performance of treatment C is easily explained.

The relatively high correlation between post-achievement test and total process posttest (0.67) and COOP test (0.70)* should be pointed out here. It may be an indication that all three tests measure a similar factor, at least to some extent. This factor may be related to content, or it may be a verbal factor or perhaps a combination of the two. As previously indicated, the achievement test was judged to have approximately 70% of its items based on higher mental process objectives whereas almost all items in the process measures were judged to be of this type. In this light the substantial correlations are not surprising.

The lack of clearcut variations in performance of the four treatment groups on the various test instruments is probably an indication of the inextricable link of <u>content</u> and <u>process</u> in science teaching, learning and testing. While it is possible to construct "pure" knowledge or content items, it is impossible or makes little sense to construct "pure" process items devoid of content. As a matter of fact, as was contended earlier in this paper, processes in scientific inquiry are general intellectual

---

*See Table VII.

processes which become specific only when linked to specific content. To this extent, the process measures may be invalid measures of process as they of necessity contain content elements.

B. General Conclusions

A number of general conclusions based on test results, non-test data or observations by the writer, participating teachers and resource personnel follow.

One of the most prominent features of this study, so far little discussed, was the inclusion of treatment group B which included classes taught the experimental curriculum by teachers who in no way participated in the preparation of these materials. The intention, of course, was to get some indication of the importance of teacher participation in the construction of a curriculum emphasizing the process dimension. The à priori feeling was that the approach taken in the prepared materials would be so novel to these teachers that their lack of understanding and conviction of the process approach would be reflected in their pupils' achievement on some of the instruments used in this study.

But even before test results became available, it became apparent that the teachers who participated in the preparation of the materials generally enjoyed no special advantage over those that did not. While the participating teachers understood the programming of the processes into the subject matter, they received little direction and help in teaching by the process approach. As it turned out, they all struggled with this aspect of the experiment and it is safe to say that even teachers in treatment A did not use one process approach in their teaching.

Unfortunately then, aside from showing little difference between treatments A and B, the test analyses did little to confirm or deny this initial assumption. Once again the major reason cited is that of the short duration of the study. Had the experimental period been longer, trends in favour of treatment A may have become noticeable.

In any case, the notion of teaching by the process approach needs much further study.

A related issue deals with the feasibility of teaching science by the approach exhibited in the experimental curriculum. Quite apart from the results of the tests, personal observations of classrooms by the writer and other members of the Project, lead to the conclusion that generally this approach to science teaching is feasible. As a matter of fact, pupil interest, participation and reaction were extremely favorable in most cases. Teachers, although feeling that they required more preparation time, were convinced of the advantages of this method of teaching science over the more traditional lecture-demonstration types. Some of the advantages mentioned were: increased pupil interest, easier acknowledgement and adaptation to individual differences, the ability to remove the memorization-oriented objectives and related student behaviours, making the student mentally and at times physically active during learning sequences, allowing students to display initiative in relation to following particular avenues while carrying out investigations, and many more. The approach does, however, require teachers to avoid giving answers and lecturing; rather, the teacher's role becomes one of

planner, facilitator, evaluator and guide. From the variety of classroom situations encountered in this study it can also be concluded that the important variable deciding whether the approach is feasible and workable is not the equipment, or lack thereof, found in a school, nor the availability of a laboratory, but a well-prepared teacher convinced of the approach and experienced in practicing it.

Since the inception of this study, the emphasis on the process dimension in science teaching has become much more widespread. This approach has been adopted in most elementary science programs on the North American Continent. With many agencies attacking the same problem from a variety of vantage points, there is some hope that more help to the practising teacher is forthcoming regarding ways and means of teaching for the processes in scientific inquiry.

The testing program which was conducted and a number of related issues provide the subject for further discussion. The ease and the effectiveness of presenting test stimuli by means of audio-visual aids cannot be emphasized enough. Their superiority over lengthy written statements of problems or descriptions of experimental procedure at the upper elementary or junior high school level is self-evident. In addition, the ease of constructing higher mental process questions based on visual stimuli on the one hand and <u>An Inventory of Processes In Scientific Inquiry</u> on the other makes the construction of this type of test a relatively simple matter; much simpler, in any case, than constructing higher mental process questions on content alone. The one disadvantage that comes to mind regarding the use of the film loop

format is that the pupil cannot refer back to written data but has to rely on his powers of observation and concentration. This problem could be alleviated if a pupil took the test individually in which case he could stop or view the loop as often as he wished.

III. RECOMMENDATIONS

For Practice in the Classroom

Even though test results did not bring out the superiority of the process approach to the teaching of junior high school science, personal observations of the writer and members of the Project of classes in action lead to the conclusion that this type of curriculum can be a workable one for process emphasis.

There are still many obstacles to be overcome and weaknesses to be removed, but the relatively successful implementation of the experimentally designed curriculum supports this conclusion.

One of the preconditions is that the student be made aware that memorization of facts, impassive listening to lectures and cramming for exams are no longer the order of the day. Rather, the approach is predicated on the idea that the student be mentally active in the learning situation rather than in the testing situation. He is constantly challenged to isolate and define problems, identify hypotheses, browse through literature, suggest designs of experiments, use his ingenuity in following designs, structure his observations, give reasons for observations, communicate his results, revise hypotheses, etc. The emphasis is clearly on student involvement, especially mental involvement.

Obviously the above are not common conditions or activities found in the typical science classroom. The best evidence to this fact is that it takes students some time to reorient themselves, especially the ones that found the "memorization-regurgitation" method easy, comfortable and so much more secure. This idea of students having certain preconceived notions about how science should be taught and learned in the junior high school (the "set" presumably being established in the elementary school) has definite implications for science instruction at the elementary level. Fortunately, most modern elementary science programs recognize the futility of teaching facts or using the lecture-demonstration approach, and have turned instead to a study and use of some basic processes in scientific inquiry.

Another factor pointed to the need for this type of elementary science background: it was found relatively difficult to teach the more complex processes in scientific inquiry such as hypotheses, model building, operational definitions, etc., without the student having some understanding of the more basic processes (observations, classification, inference and prediction, etc.). In this regard the hierarchy of processes proposed by Gagné in Science - A Process Approach seems quite plausible. Learning science in the proposed manner in elementary school would lessen or remove the set that students entering junior high school now typically possess.

If this type of science instruction is difficult for the student to orient to, it is apparently more difficult for the teacher. Teachers have long been used to being in the limelight while conducting their classes. In the process approach the spotlight is on the learner; the teacher only "interferes" in isolated instances. There is a busy atmosphere in the

classroom; desks may not be in neat tidy rows, some students may be reading, others experimenting, and still others talking or writing on the board. The teacher may be questioning or answering one student or a small group of students, he may be obtaining a particular piece of apparatus for another group, or he may be silently wandering about the room making more or less subjective judgments about how particular students are performing in terms of some particular process objective. The time spent in lengthy preparation for lectures, where every detail of content is correct and well understood by the teacher, is now spent in making sure materials and resources are available for class, holding student conferences, finding audio-visual aids to support class activities and generally planning for and anticipating problems.

Unfortunately, a teacher's preparation time for any one class is limited, and preparing for an activity oriented class or set of classes is much more time consuming than preparing for lectures. The curriculum development approach developed by the Portland Public School System is one step in the right direction toward alleviating this difficulty. Their CYBEX (48) system is predicated on the involvement of all science teachers in their system for the building of a science curriculum. In this way, a bank of activities and approaches is set up for a variety of topics and content areas. Teachers are responsible for curriculum development and the implementation and constant revision of their science curriculum.

The notion of conscientiously integrating as many processes as possible into any one investigation seems a workable one as does the idea

of allowing certain processes to recur from time to time. The fact that teachers and students will encounter a process periodically means that not only should a student come to know what a particular process is and means, but it also requires that a student grow in his understanding of a particular process in scientific inquiry. This in turn implies that the teacher have a fairly comprehensive understanding of all the processes in scientific inquiry. Unfortunately, this knowledge and understanding usually cannot be completely gleaned from books, and therefore the necessity for intensive workshops and inservice seminars is foreseen.

The recurrent inclusion of certain processes into the experimental curriculum was based on another assumption: if students encounter a certain process often enough and in a variety of contexts, then sooner or later the amount of guidance connected with the teaching of any one process can be reduced and at times completely omitted. In practice this idea seemed to hold, especially for the simpler processes, while the teaching of the more complex processes usually requires far more guidance and a larger number of encounters.

For Evaluation of the Process Approach

There are two immediate implications arising from the testing program conducted in this study: one, the inadequacy and relative fruitlessness of "end of the term" testing, and second, the futility of testing for gain in process objectives through the use of "end of the term" - type measures if the term is only of two months duration.

It is apparent that evaluation of progress in process skills should be an ongoing and continuous enterprise even if this necessitates individual testing and the use of more or less subjective instruments or schemes. In this regard, the evaluation scheme proposed by Science - A Process Approach has much to recommend it. Most computer-assisted instructional systems also recognize the need for small-step progress and immediate evaluation and feedback. The student's record book may prove to be a valuable source upon which some evaluation might be based.

The second implication arises from the fact that most of the conclusions based on tests showed relatively little difference between the two experimental groups (A and B) and treatment C. The period of instruction will need to be much longer before significant differences using "end of term" tests are likely to be detected. A full year period or more seems much more reasonable.

The ability to detect significant differences leads naturally to the tests used, particularly the process measures. It is quite likely, although there is no direct evidence for this, that the content, at least, of the process measures will have to be refined or modified. At the same time, the idea of using various levels of testing based on the "hierarchy of meaningfulness" still seems quite sound. This argument once again brings out the inextricable link between content and process in science teaching, and the improbability of constructing "pure" process measures. Nevertheless, constructing test items based on An Inventory of Processes in Scientific Inquiry, when this scheme was used in the teaching and learning situation, still seems appropriate; much more so in fact, than using Bloom's taxonomy in trying to construct higher mental process items.

There is no doubt, however, that the validity of this type of measure has to be more clearly established. In some way tests of this type have to be analyzed (perhaps factor analysis can help here) in order to determine exactly what variables are being measured. Carefully controlled studies will have to be conducted removing, if possible, such confounding factors as reading comprehension, science content and previous experience.

The use of audio-visual aids in testing was shown to be quite workable and useful in partially removing a confounding reading or verbal factor. If end of the term testing is desirable, this means of presenting stimuli and perhaps even for recording responses shows much promise. This may be particularly true of the lower levels of schooling where written tests may measure little besides reading comprehension. As already indicated, this format may be easily adapted to individualized testing.

A connected recommendation deals with the inadvisability of administering a battery of tests in a short time period. Even though pre and posttesting was spaced over five or six consecutive science classes, pupils tended to become uninterested, negligent, tired, and somewhat apathetic toward the end of the testing sequence. This was particularly true for the posttests which were administered during early June, a period when students tend to be particularly active. It is recommended here that test batteries be spaced and administered well before the end of the year or well before other holiday periods.

For Further Research

(a) This study attempted the development and evaluation of a process approach to the teaching of science restricting itself solely to

the cognitive domain of science teaching objectives. There is already some evidence that students instructed by this approach show significant gains in the affective* and psychomotor domains. More instruments for assessment of growth in these areas and more theoretical scrutinizing need to be done in this regard.

(b) Little is known about whether the process approach is suited to a particular type of student in terms of ability, interest or aptitude. Nor is there evidence to suggest that some students benefit more from formal instruction than the proposed process approach. Investigation of this area could prove very crucial to future curriculum development.

(c) Aside from recurring encounters with various processes, far too little is known about how to teach certain processes such as hypotheses, model-making, etc. Action research to try various means and techniques will have to be attempted. A start in this direction has already been made by some members of the Project. Pupil note books and teachers' anecdotal records were examined after the experiment in order to determine if they contained any clues as to how some of these processes might be taught.

(d) There is a definite need for longitudinal studies to determine long-range effects of this type of instruction.

(e) One area that may lead to interesting research is that of testing treatment groups exposed to curriculum materials for equal time

---

*Dr. M. A. Nay and R. K. Crocker have recently written a very comprehensive article entitled "Science Teaching and the Affective Attributes of Scientists" (unpublished) which explores the Affective domain of science teaching objectives.

periods versus exposing groups to equal content areas. Typically, and this proved to be so in this investigation, traditional treatments tend to "cover" content areas much faster than activity-oriented treatments leading to more content being covered in traditional classes during equal time periods. This factor may give traditional treatments an advantage especially on achievement tests, and the whole area is worthy of further research.

(f) There may be merit in controlling the teacher variable in a similar fashion to the control introduced by Klopfer (35) in his validation of the TOUS. He tested prospective teachers as to their understanding of the scientific enterprise and then added this dimension as a stratification factor. This control may prove to be more effective than using an additional treatment group as defined in this study.

(g) Very little attention was paid to the types of facilities and science rooms available to each treatment and classes within treatments. The effect of this factor should be researched as typically it is assumed that better equipped schools lead to better learning in science.

(h) More instruments measuring process objectives with greater statistical confidence are badly needed. In this regard, the effectiveness and feasibility of using audio-visual aids as resources need to be further explored. Adapting test instruments to the level of schooling is another area not studied very thoroughly. Only very recently have pictorial tests based on other than "word" stimuli been constructed. A radical departure from typical paper-and-pencil tests may be necessary. In addition, the

whole area of continuous evaluation techniques deserves much more attention than it has received to date.

(i) As indicated previously, this investigation is one of the first really concerned with process teaching in this manner. There are other means of accomplishing similar objectives such as the approach taken in Science - A Process Approach. Comparative studies investigating and evaluating respective gains might be fruitful by giving evidence of the possible superiority of one method over another for certain process objectives. Only longitudinal comparative studies show any promise in this area.

LIST OF REFERENCES

LIST OF REFERENCES

1. American Association for the Advancement of Science, Science - A Process Approach. New York: Xerox Corporation, 1968.

2. Association for Supervision and Curriculum Development, The Way Teaching Is, Report of the Seminar on Teaching. Wash., D.C.: National Education Association, 1966.

3. Atkin, J. Myron, "Science Education, Comments on Two Approaches," (Toronto: Ginn and Company, 1966), Pp. 1.

4. Ausubel, David P., "An Evaluation of the Conceptual Schemes Approach to Science Curriculum Development," Journal of Research in Science Teaching, (3:11; 1965).

5. Biological Sciences Curriculum Study, High School Biology. Chicago: Rand McNally & Company, 1967.

6. Biological Science Curriculum Study, Processes of Science Test, Form A. New York: The Psychological Corporation, 1962.

7. Bloom, Benjamin S., et al., Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I, Cognitive Domain. New York: Longmans, Green & Co., Inc., 1956.

8. Brandwein, Paul F., Elements in a Strategy for Teaching Science in the Elementary School, The Burton Lecture. New York: Harcourt, Brace & World, Inc., 1962.

9. Bruner, Jerome S., The Process of Education. Cambridge, Mass.: Harvard University Press, 1961.

10. Campbell, Donald T. and Stanley, Julian C., "Experimental and Quasi - Experimental Designs for Research on Teaching", Handbook of Research on Teaching, edited by N. L. Gage. Chicago: Rand McNally, 1963. Pp. 578.

11. Chemical Bond Approach Project, Chemical Systems. New York: Webster Division, McGraw-Hill Book Company, 1964.

12. Chemical Education Material Study, Chemistry - An Experimental Science. San Francisco: W. H. Freeman and Company, 1963.

13. Ibid., pp. vii - viii.

14. Commission on Science Education, American Association for the Advancement of Science and Science Teaching Center, University of Maryland, Sixth Report of the International Clearinghouse on Science and Mathematics Curricular Development, 1968, edited by J. David Lockard. College Park, Md.: University of Maryland, 1968.

15. Conant, J. B., Harvard Case Histories in Experimental Science. Cambridge, Mass.: Harvard University Press, 1957.

16. Cooley, William W. and Klopfer, Leopold E., Test On Understanding Science. Princeton, N.J.: Educational Testing Service, 1961.

17. Dewey, John, Democracy and Education. New York: MacMillan Company, 1961, paperback edition.

18. Earth Science Curriculum Project, Investigating the Earth. Boston: Houghton Mifflin Company, 1963.

19. Educational Testing Service, Cooperative Testing Division, Cooperative Science Tests Handbook. Princeton, N. J.: Educational Testing Service, Inc., 1964.

20. Ibid., p. 57

21. Ibid., p. 62.

22. Ibid., p. 63.

23. Feldt, Leonard S., "A Comparison of the Precision of Three Experimental Designs Employing a Concomitant Variable", Psychometrika, (December, 1958), Pp. 335-353.

24. Gagné, Robert M., The Conditions of Learning. New York: Columbia University Press, 1960.

25. Gagné, Robert M., "Psychological Issues in Science - A Process Approach", The Psychological Bases of Science - A Process Approach, by the AAAS Commission on Science Education. Wash., D.C.: American Association for the Advancement of Science, 1965. Pp. 4.

26. Ibid., p. 4 - 5.

27. Ibid., p. 5.

28. Harvard Project Physics, Project Physics, interim version. New York: Holt, Rinehart & Winston, Inc., 1969.

29. Hukins, Austin A., "A Factorial Investigation of Measures of Achievement of Objectives in Science Teaching", (Unpublished Ph.D. dissertation, University of Alberta, 1963), Pp. 103.

30. Hunt, T. W., Andrews, H. C. and Hedges, H.C., Science Activities. Toronto: W. J. Gage Ltd., 1959.

31. Hutchinson, John, "Automated Science Curriculum: An Experimental Science Program", (Unpublished Ph.D. dissertation, Oregon State University, Corvallis, 1965).

32. Hutchinson, John, "Science Teacher's Adaptable Curriculum", Journal of Research in Science Teaching, (December, 1964), Pp. 356 - 363.

33. Ingebo, George, Report of the International Clearinghouse on Science and Mathematics Curricular Developments, 1966, "Teacher's Automated Guide (TAG)", Pp. 230 - 231.

34. Ivany, George W., "A Comparison of Expository and Hypothetical Modes of Teaching Science", (Unpublished Ph.D. dissertation, University of Alberta, 1965), Pp. 41.

35. Klopfer, L. E. and Cooley, W. W., "The History of Science Cases for High Schools in the Development of Student Understanding of Science and Scientists", Journal of Research in Science Teaching, (No. 1, 1963) Pp. 33 - 47.

36. Ibid., p. 33.

37. Lizonbee, Lorenzo, "Testing What For?", The Science Teacher, (1966), 33 (5), Pp. 27 - 29.

38. National Science Teachers' Association, Theory Into Action...in Science Curriculum Development. Wash., D.C.: NSTA Curriculum Committee, 1964. Pp. 17 - 18.

39. National Society for the Study of Education, Rethinking Science Education, edited by Nelson B. Henry, Part I. Chicago: University of Chicago Press, 1960. Pp. 34 - 37.

40. National Society for the Study of Education, Science Education in American Schools, Forty-Sixth Yearbook, Part I., Chicago: University of Chicago Press, 1947.

41. Nay, Marshall A., "The Discovery Method of Teaching", The ATA Magazine, (Vol. 48, No. 5, 1968), Pp. 20 - 23.

42. Nedelsky, Leo, Science Teaching and Testing. New York: Harcourt Brace and World, Inc., 1965.

43. Obourn, Ellsworth S., An Analysis and Checklist on the Problem-Solving Objective, Circular #481, Office of Education. Washington, D.C.: U.S. Department of Health, Education, and Welfare, June, 1956.

44. Parker, Cecil J. and Rubin, Louis J., Process as Content: Curriculum Design and the Application of Knowledge. Chicago: Rand McNally Co., 1966. Pp. 48.

45. Ibid., p. 19.

46. Ibid., p. 44.

47. Physical Science Study Committee, Physics. Boston: D. C. Heath and Company, 1965.

48. Portland Public School System, CYBEX and TAG, Approaches to Curriculum Development. Portland, Oregon: Portland Public Schools.

49. Portland Public Schools Testing Service, Portland Science Test. Portland: Portland Public Schools, 1963.

50. Portland Science Test Committee, Portland Science Test. Portland: Portland Public Schools, 1963.

51. Psychological Corporation, The, Information Leaflet. New York: The Psychological Corporation, 1967.

52. Schalock, Henry D., Beaird, James H. and Simmons, Helen, "Motion Pictures as Test Stimuli: An Application of New Media to the Prediction of Complex Behavior", Final Report, Title VII Project #971. U.S. Office of Education, 1964.

53. Schwab, Joseph J., "The Structure of the Disciplines: Meanings and Significances", The Structure of Knowledge and the Curriculum, edited by G. W. Ford and L. Pugno. Chicago; Rand McNally and Company, 1964.

54. Schwab, Joseph J., "The Teaching of Science as Enquiry", The Teaching of Science, Joseph J. Schwab and Paul F. Brandwein. The Inglis and Burton Lectures, Harvard University, 1961. Cambridge, Mass.: Harvard University Press, 1962.

55. Ibid., Pp. 55 - 56.

56. Ibid., Pp. 65 - 66.

57. Ibid., p. 67.

58. Stotler, Donald, "I believe in the Modern General Science Approach", NEA Journal, (October, 1963), Pp. 55 - 56.

59. Stotler, Donald, "Science, Heuristics and Humanism", The Science Teacher, (October, 1965).

60. Stotler, Donald and Hutchinson, John, Third Report of the Information Clearinghouse on New Science and Mathematics Curricula, (March, 1965), "The STAC Approach to Curriculum Planning Using Keysort Cards", Pp. 110 112.

61. Suchman, J. R., The Elementary School Training Program in Scientific Inquiry, Title VII, Project #216, National Defence Education Act, Grant No. 7-11-038. Urbana, Ill.: University of Illinois, 1962.

62. Suchman, J. Richard, Inquiry Development Program. Chicago: Science Research Associates, Inc., 1966.

63. Taba, Hilda, Curriculum Development. New York: Harcourt Brace and World, Inc. 1962.

64. Watson, Fletcher, G., "Research on Teaching Science", Handbook of Research on Teaching. N. L. Gage, Editor. Chicago: Rand McNally, 1963. Pp. 1054.

65. Welch, Wayne W., and Walberg, Herbert J., "A Design for Curriculum Evaluation", Science Education, (Vol. 52, No. 1, 1968), p. 12.

66. Winer, B. J., Statistical Principles in Experimental Designs. New York: McGraw-Hill Book Company, 1962. Pp. 578.

67. Ibid., p. 101.

68. Ibid., p. 579.

69. Ibid., p. 579.

70. Ibid., p. 579.

71. Ibid., p. 586.

72. Ibid., p. 586

73. Ibid., p. 620.

74. Ibid., p. 620.

75. Ibid., p. 621.

APPENDICES

APPENDIX A

AN INVENTORY OF PROCESSES IN SCIENTIFIC INQUIRY

(Fifth Draft - August, 1968)

Prepared by the Edmonton Junior High Science Project

M. Nay, R. Melnychuk, E. Mokosch, D. Powley,
R. Roberts, D. Thomas, H. Visscher, D. Whitlock

A. GENERAL PROCESSES

The following are common to all the processes in scientific inquiry listed below:

1. Communication: all of the basic communications skills (reading, writing, speaking, listening, etc.)

2. Reflective, logical thought: analysis, synthesis, deduction, induction, etc.

3. Evaluative, critical thought: checking reliability and validity, identifying inconsistencies and errors, etc.

4. Creativity: inspiration, intuition, imagination, innovative thinking, originality, etc.

B. SPECIFIC PROCESSES

I. PREPARATION

1. Identifying and formulating a problem

   (a) speculating about a phenomenon

   (b) identifying variables

   (c) noting and making assumptions

   (d) delimiting the problem

2. Seeking background information

   (a) recalling relevant knowledge and experience

   (b) doing literature research

   (c) consulting people

3. Predicting

4. Hypothesizing

5. Design for collection of data through field work and/or experimentation

   (a) defining the independent and control variables

   (b) defining the procedure and sequencing the steps

   (c) identifying needed equipment, materials and techniques

   (d) recognizing safety precautions

   (e) devising method for recording data

II. <u>COLLECTION OF DATA</u>

6. Procedure

   (a) collecting, constructing and setting up the apparatus or equipment

   (b) doing field work and/or performing the experiment

   (c) identifying the limitations of the design (as a result of failures, blind alleys, etc.) and modifying the procedure (often by trial-and-error)

   (d) repeating the experiment (for reproducibility, to overcome limitations of initial design, etc.)

7. Observing and observations

   (a) qualitative (using senses, etc.)

   (b) quantitative (measurements, readings, calibrations, counts of objects or events, estimations, approximations, etc.)

(c) specimens

(d) graphical (charts, photographs, films, etc.)

(e) unexpected or accidental observations (serendipity)

(f) noting precision, and possible errors and mistakes

(g) recording data (describing, tabulating, diagramming, photographing, etc.)

III. PROCESSING OF DATA

8. Organizing the data

(a) ordering

(b) classifying

(c) comparing

(d) other methods

9. Representing the data graphically

(a) drawing graphs, charts, maps, diagrams, etc.

(b) interpolating, extrapolating, etc.

10. Treating the data mathematically

(a) computing (calculating)

(b) using statistics

(c) determining the precision and accuracy

IV. CONCEPTUALIZATION OF DATA

11. Interpreting the data (interpretations and inferences)

(a) suggesting explanations and/or deriving generalizations or "conclusions"

(b) assessing validity of initial assumptions and hypotheses

12. Formulating operational definitions

13. Deriving mathematical relationships

14. Developing a theory (building "mental models")

V. <u>OPEN-ENDEDNESS</u>

15. Seeking further evidence to

   (a) increase the level of confidence in the explanation or generalization

   (b) test the range of applicability of the explanation or generalization

16. Identifying new problems for investigation because of

   (a) the need to study the effect of a new variable

   (b) anomalous behaviour or unexpected observations

   (c) incompleteness ("gaps") and inconsistency in the theory

17. Applying the discovered knowledge

EXPLANATION OF THE PROCESSES IN SCIENTIFIC INQUIRY

(Based on the Fifth Draft of the Inventory)

A careful look at the Inventory will show that scientists work mainly with their brains rather than their hands. Yet the "test of Truth" is derived from the observation (data) obtained from field work or experiments in the laboratory. Therefore, scientific activity requires an appropriate combination of thought and physical work.

It should be pointed out that although the processes are numbered from one to seventeen, no rigid order is implied. Scientific investigations do not proceed always in an orderly manner from step one to seventeen. Furthermore, in any given investigation it may not be necessary to include all the steps. Finally, two or more processes often overlap or are "telescoped" into a single operation.

I. PREPARATION

This major division includes activities encountered before the collection of data begins.

1. Problem

There are many things and events in the universe about which every person wonders and speculates. Thus, accidentally or deliberately we are continually identifying problems, any one of which could be investigated to obtain an answer or solution. For instance a problem may arise from observing a phenomenon (e.g., a rainbow), noting discrepancies in nature (e.g., a liquid flowing upward), or merely from curiosity. The problem may be stated in the form of a question:

Example: What is a rainbow?

or as an infinitive phrase:

Example: To determine how a liquid can flow upward.

The investigation can be more fruitful if the problem is clearly defined or formulated. We need to identify those variables which will suggest the kind of data to collect through field work and/or experimentation in the laboratory, and how to collect it. Variables are factors, conditions or properties which operate within or influence the phenomenon under investigation.

Example of variables: weight, temperature, distances, water supply, cosmic radiation, animal population density, etc.

Often to be able to find an answer to a problem, we have to make assumptions.

Examples of assumptions: certain variables have no effect on the phenomenon, reliable data can be obtained, etc.

Finally, the problem may have to be dilimited. Otherwise too big an area may have to be covered, making the experiment too difficult to perform, or the data too unreliable or confusing to be of much value. This immediately places a limit on the knowledge we can obtain about the problem. However, what we do find out will probably be more reliable and valid.

2. Background Information

Before any problem can be investigated or solved, or perhaps even before it can clearly be defined, some information is needed. This consists of background theory, knowledge of what other scientists have done on the same (or similar) problem, ideas on the apparatus and techniques to use, precautions, etc. The investigator may already have some useful knowledge for solving the problem. Usually, he can get a great amount of information through literature research; that is, by reading appropriate books, papers, articles, etc. Listening to speeches and discussing the problem with other people is another method of obtaining the necessary background. (In school, the people consulted will be mainly the science teachers and classmates, although occasionally it might be possible to consult a scientist who is an expert in the area being investigated.)

3. Predictions

Predictions are based on the fact that the universe operates in a regular manner, and that we expect known phenomena to occur always in the same way.

Examples: the movements of the moon, changes in season, effect of gravity, flow of heat, boiling, etc.

This behaviour of nature allows us to make predictions based on past experience. Whether a prediction is right or wrong depends on the probability of something happening in the expected way. For instance, our prediction that the sun will rise tomorrow morning is almost certain to be correct, since the probability of the sun rising is near certainty. The prediction of "heads" in coin-tossing will be right only fifty per cent of the time. Predictions for rain in a specified place at a

definite time will probably be wrong most of the time, because the probability of this happening is very low.

A prediction may or may not suggest specific observations to make or procedures to follow to test the prediction.

4. Hypothesis

A hypothesis is a possible or tentative explanation for a phenomenon. This may be anything from a conjecture, guess or assumption to an explanation which is highly probable in the light of known observations or facts.

Examples: The hypothesis that there are other planets in the universe with thinking beings inhabiting them is not much more than a (statistical) guess. On the other hand, the existence of Pluto and Neptune was hypothesized to account for some anomalous astronomical behaviour.

It is usual to try to test (prove or disprove) the validity or truth of a hypothesis by collecting more data. To do this a working hypothesis might be stated to guide the investigation.

Example:

Hypothesis: Birds migrate south in the fall because of the decreasing length of daylight.

This suggests an experiment in which two variables may be studied for birds in a windowless case: length of artificial daylight and the direction of flight.

The following example illustrates the distinction among problem, prediction and hypothesis:

Situation: If a glass of water is placed on a paper towel which is then whisked out from underneath....

Problem: Will the glass of water remain upright?

Prediction: Probably not since experience with imbalances of forces suggests that in this case the glass will move and probably tip over.

Hypothesis: Involved is not only an imbalance of forces, but also inertia. Consequently, if the towel is whisked out from under the glass with a quick, jerky motion and the towel is dry, the effect of inertia will be dominant. The glass will remain upright.

5. Design for Collection of Data

A solution or explanation for a problem may be stated first as a hypothesis, but the truth or validity thereof can be established only on the basis of the right observations or data. This data may come entirely from experiments performed in the laboratory (as in chemistry or physics), or from field work which often is combined with laboratory activity (as in geology and biology).

The value of planning the method or procedure for collecting the data is obvious. First the researcher has to deal with the variables. It is common practice to study only some of the variable (called independent variables) while keeping all the others under control ("held constant").

Example: Referring to the above experiment with a glass of water on a towel, if the variable being studied is the speed of removing the towel, then the controlled variable are the amount of water in the glass (same amount in all the experiments), wetness of towel (keep it dry), roughness of table, etc.

After the effect of changing one variable has been studied, it can then be controlled in turn while the effect of changing some other variable is studied (e.g., varying the amount of water in the glass, from zero to full). This aspect is referred to again under "open-endedness" in Process 16 (b).

Having made a decision about the variables, the scientist then develops a method of collecting the data including the apparatus needed, the steps to be followed, the precautions to be taken, and manner in which the data is to be recorded (e.g., narrative and tables in a notebook, charts from electronic measuring instruments, photographs, specimens, etc.).

II. COLLECTION OF DATA

This general category includes the activities associated with the actual collection of data, in the field and/or laboratory. (In science teaching this phase is commonly referred to as "experimentation" or "laboratory work". This designation is incomplete since in many sciences much data is obtained in activities which do not resemble experiments, and in locations outside of a room called a "laboratory".)

6. Procedure

The scientist starts his data collection by following the design. He may be fortunate in collecting sufficient reliable and reproducible data without undue deviation from the design. However, it is more probable that he will encounter difficulties of varying gravity. He may have trouble getting or building the required equipment, and setting it up. The recording system may not be sufficiently sensitive or stable. He may have to learn by trial-and-error how to perform the experiment properly. He performs the experiments, but the results do not make any sense (a "blind alley"), or give him only limited information. He tries a procedure repeatedly but without success (a "dead-end"). In the field he may search different areas for specimens, but may find few or none. These and countless other kinds of difficulties may result in failure. If he accepts the failure as temporary, he may do more reading, thinking and designing - and then try again. However, if he cannot see a way out,he will drop the problem under investigation and go on to another one.

All of the above obstacles (some foreseen and others not) to success are part of the limitation of the design. Throughout the entire operation the scientist records for future reference, everything he does, whether successful or not.

7. Observation

The observations made constitute the data of research, and many of them become part of the theory as facts.

The scientist develops keen power of observation. He tries to be systematic so that nothing goes un-noticed in his investigation. As a result he is prepared for the accidental or unexpected events when they occur (many of which give rise to more important problems then the original one). The data is of various kinds and appears in various forms as indicated in process 7 (a), (c), and (d), of the inventory.

The scientist also takesinto account the precision and accuracy of his results (including possible erros, significant figures, etc.). The precision may be high in that every time he does a given experiment he gets the same result (within experimental error). However, the result itself may be inaccurate because the apparatus was not working properly, there were impurities in the chemicals, etc.

Analogy: In target practice, the precision is high if the bullet holes are all close together. For the accuracy to be high, these holes should be near or at the bull's eye.

III. PROCESSING OF DATA

This general category deals with what happens to the data after collection, but before interpretation. Putting the data in some more systematic form may reveal generalizations more readily, help to identify mathematical relationships between variables, etc. Actually the strict separation between data processing and interpretation is a little artificial, as some conceptualization usually has already taken place when the data is being processed.

8. Organization of Data

To make data more meaningful, it often has to be rearranged, compared, ordered or classified. Sometimes this step is performed before the actual performance of an experiment, as is the case when suitable variables are tabulated before making the actual observations. Usually, however, data is recorded in some "rough" fashion and has to be organized in some more compact and meaningful way.

9. Graphical Representation

This process involves systematizing the data graphically, in order to obtain "hidden" information and thus facilitate interpretation and make it more complete. Such "hidden" information is found in extrapolations and interpolations, intercepts, contours, cross-sections, and isopleths (e.g., isobars, isotherms, isoclines, etc.). These graphical representations often make possible predictions (e.g., rainfall from weather maps, oil deposits from geological cross-sections, etc.).

10. Mathematical Treatment

Quantitative data is often processed mathematically for a number of reasons. One is to make it more meaningful by deriving constants from multiplication or division of data, rates of change, slopes and areas from graphs; determining statistical distributions; etc. Another is to help define the "level of confidence" that can be placed in the observations through such operations as averaging, finding per cent errors, etc. In the mathematical treatment, often the chance effects of uncontrolled or unidentified variables become evident.

IV. CONCEPTUALIZATION OF DATA

This general category deals with the process of bringing conceptual understanding and order into the facts (observations or data). Here we are concerned with the PRODUCT or CONTENT of a scientific discipline. The relevant information identified under the process of "background knowledge" is vital for this stage of scientific inquiry.

11. <u>Interpretation of Data</u>

(In science teaching, this process has been seen mainly as one of "drawing a conclusion." The word "conclusion" is misleading since, although a given investigation may be concluded, it gives rise to a host of new problems for investigation.)

A scientific investigation is undertaken to find an explanation or solution to a problem. In due course a body of observations is obtained from which generalizations (laws, principles) are deduced, and/or for which appropriate explanations are devised. Of primary concern are composition and structure (e.g., living tissue), function, relationships (eg., pressure versus temperature for gases), and cause-and-effect (eg., heating water). Interpretation involves <u>inference</u> since often the mind must go "beyond the facts" for an appropriate generalization or explanation.

It is essential to note that it is not easy to answer even a simple problem unambiguously by experiment or field work. Therefore, vigilence must be exercised in interpreting data; that is, no more should be inferred than is justified. There is always a limit on the new knowledge that can be obtained in any single investigation.

At this stage the validity of initial assumptions and hypotheses must be assessed. New problems have arisen in the course of the investigation (openendedness). If these stem from errors and weaknesses in the experimental design, then the initial hypotheses and assumptions may be reformulated in preparation for a repeat of the investigation.

12. <u>Operational Definitions</u>

Formulating operational definitions is only indirectly a process in scientific inquiry; it is primarily an aspect of the language of science. Scientists find it useful and convenient to use a word or a brief phrase to identify the operation of an object or an event in nature. Therefore, an operational definition encompasses the minimum description or minimum action needed to identify an object or event. Often more than one definition is possible, the choice depending on suitability.

Example: Define the word "boiling".

1. Boiling is the event in which bubbles are formed continually throughout a heated liquid, and rise to the surface where they "break."

2. Boiling is the event in which the maximum constant temperature (as measured with a thermometer) is obtained for a heated liquid.

13. Mathematical Relationships

Usually this process is part of the interpretation of data. One reason why it is shown separately is that considerable importance is attached to quantitative generalizations in science. An important outcome is that manipulation of a mathematical relationship may often provide information which was not apparent from the data. Another important reason is that scientific relationships can be derived entirely mathematically (e.g., Einstein's equation relating matter and energy, $E = mc^2$). In both situations new problems may be identified for investigation, to prove or disprove the ideas deduced mathematically.

14. Development of a Theory

Whereas interpretation is a process concerned with the data from the investigation, theory building involves fitting and integrating the new knowledge into the theory already existing; or revising this theory to accommodate both the new and old facts. Facts, laws, principles, concepts, conceptual schemes, hypotheses, etc. are the "bones and flesh" of a theory. Considerable insight and intuition (involving going "beyond the facts") are called for to bring about a consistent integration of these elements of a theory. Since much of this theory goes beyond what can be observed directly (e.g., atoms, molecules, genes,etc.), and exists primarily in the scientist's mind, the term "mental model" building has become popular in science (and science teaching).

It is at this stage of inquiry that the provisional (tentative) and self-correcting (revisionary) nature of scientific knowledge is most apparent. For instance, the new findings may show some of the old principles to be erroneous or obsolete and in need of revision. Or the new data may help to settle which of two existing theories is more plausible. In any event, a host of new problems stem from gaps, inadequacies, conflicts, and inconsistencies in the theory, and new assumptions and hypotheses may be formulated as a prelude to new investigations.

V. OPEN-ENDEDNESS

Science is always an unfinished business. One often hears that each investigation raises many more questions than it answers. This aspect of the scientific enterprise is often referred to as "open-endedness." In this inventory, the open-endedness is seen in terms of the need for further support for a generalization or explanation, the application of this generalization or explanation in explaining new phenomena (some of which have been long observed but not understood), and the pursuit of new problems related to the one being investigated.

15. Further Evidence

Obviously, the level of confidence in a generalization or explanation will be low if errors or weaknesses in some stage of the investigation have been spotted. Those which appeared early might be remedied at the data-collecting stage, while those appearing later may require the investigation to be redone. Often there is a "pilot" stage in an investigation for the purpose of testing the design for data collection, and revealing some of the difficulties and pitfalls.

In some cases it is useful to repeat an investigation (in the same or a modified form) to accumulate more evidence for interpretation. For example, the average of several temperatures of boiling water is more reliable than a single reading. Or it may be desirable to determine a more accurate boiling point using a more sensitive thermometer and better experimental technique.

When a scientist obtains a generalization or explanation for one substance, he may want to know how other substances behave when the same variables are studied. Thus he increases the range of application of the generalization. For example, when he studies the effect of heat on several different liquids, he finds that all of them can be made to boil.

16. New Problems for Investigation

As was already mentioned under Process 14, "gaps", inconsistencies, discrepancies, or weaknesses in the theory point the way to new investigations.

One characteristic of new problem is the includion for study of at least one new variable. A standard way in which this is done in research is to make one of the control variables in a previous investigation an independent variable (see example under Process 5 above).

Anomalous behaviour or unexpected observations usually are the result of some variables not being recognized initially (hence no provision being made to control them). Often these give rise to investigations which are much more important than the one in which they appeared. For example, it is found that all liquids boil, but the boiling points differ. In addition, the boiling temperature will vary with altitude. Fleming's discovery of the effect of penicillin is a more dramatic example of the importance of "accidental" or "chance" discovery in the development of science.

17. <u>Application</u>

The scientific knowledge gained from an investigation may be applied in a number of ways. Perhaps the simplest is its use in solving some related problem arising out of the openendedness of the investigation Often knowledge gained in one area may help increase the knowledge and understanding in another. Such applications may vary in magnitude, ranging from explanations of simple concrete phenomena to providing the basis for entire scientific fields (e.g., physics in biophysics, chemistry in geochemistry). Another important application of scientific knowledge is in <u>technology</u> (applied science) through which man gains control over parts of his environment and improves his own well-being.

Investigation A - 2 (Student Sheets)

PREPARATION

Problem

To obtain the smallest particle or unit from which sugar is made.

Hypothesis

Design for Collection of Data

At the end of Investigation A - 1, you and your classmates were asked to suggest how your smallest particle of sugar could be broken down some more. Maybe it is possible to break it down even into the smallest units from which it is made. Many interesting suggestions were given. Here is a method you are to try:

1. Place a sugar cube in a 150 ml. beaker containing about 25 ml. of water. Do not stir.

   (a) Observe what happens to the sugar cube.
   (b) Stir the solution and test for sugar in the water by tasting.
   (c) Test the solution for sugar by the "cane sugar test" as follows:

   To a small quantity of cane sugar solution add one to two cc of the cobalt nitrate solution. Then add a small quantity of NaOH solution (CAUTION). A violet colour reaction is a positive test for cane sugar. The above substances result in a light orange colour and a precipitate.

2. As a "control", repeat the two tests using only tap water. (Do you get the same results?)

3. Take a filter paper and look at it under a microscope.

4. Place a folded filter paper in a funnel supported by a test tube, then place sugar grains on the filter paper. Pour 10 ml. of water onto the filter paper. Make tests on the <u>filtrate</u> for sugar.

Questions:

1. What is the meaning of the following words: filtrate, solvent, solute, solution?

2. Why was a "control" experiment included in the design?

COLLECTION OF DATA

Procedure

Follow the above design, and record any additions or changes which were not foreseen in the design.

Observations

CONCEPTUALIZATION OF DATA

Interpretation of Data

1. How did the sugar get from the filter paper into the solution?

2. Did you obtain the smallest possible particle of sugar in this experiment? Explain.

3. Was the "hypothesis" for Investigation 1 and 2 confirmed (found to be correct)? Explain.

4. In what way is the sugar in the cube different from the sugar in the water?

Operational Definitions

1. A molecule is the smallest particle or unit of which any substance is built.

2. Define "dissolving" in terms of what happens to the molecules.

Development of a Theory

On the basis of the evidence obtained in this investigation we are still not very sure that sugar crystals are made up of very tiny indivisible particles called molecules. However, scientists have obtained much evidence to support this hypothesis, and have applied the idea to all substances. They also believe that all the molecules in any given pure substance are alike. These ideas about the molecule nature of matter are an important part of the "kinetic-molecular theory".

OPEN-ENDEDNESS

Further Evidence

Suggest other evidence which might help us become more certain that all substances are composed of molecules.

New Problems for Investigation

1. List two new problems which have occurred to you while doing this investigation. What are some of the new variables in each of these problems?

2. Do library work and prepare a short report on what scientists have found out about the size of molecules.

(c) Prepare the 5% cobalt nitrate solution by dissolving 5 grams of cobalt nitrate in 100 ml distilled water. Also prepare the sodium hydroxide solution by dissolving 50 grams of sodium hydroxide in 100 ml of distilled water.

(d) Demonstrate the technique of folding and using filter paper.

(c) Discuss meaning of the words in Question 1 and have students identify each in this experiment in this investigation.

6. Procedure

Suggestions for the Teacher

(a) After the children have done the experiment, the teacher should discuss and summarize the additions and changes required in the initial design, and should make an evaluation of the design.

7. Observations

The results here are qualitative. There are two parts to this experiment. The purpose of the first is to identify the dissolving property of and tests for sugar. In the second part, students "observe" the sugar disappear through the filter paper by dissolving, and appear in the solution (as confirmed by positive tests; sweet taste and a violet colour). The tests for the "control" solution are negative (no sweet taste and no violet colour).

Suggestions for the Teacher

(a) Review "observing," "observations," and "recording." Stress the importance of careful observation and recording only that which is observed.

(b) In class discussion, summarize the observations made by students. Each student should place a line under his own observations before taking down any observations which have come up in the discussion and which he missed.

Investigation A - 2 (Teacher's Guide)

Concepts

1. Matter is composed of very small basic particles or units called molecules.

Processes Involved

1. Problem

   In a sense this investigation is a continuation of Investigation A - 1. The problem is the same, but the design is different since a dead-end was reached with the previous one (see page T - 4).

2. Background Information

   This was done under "openendedness" in Investigation A - 1/

4. Hypothesis

   This remains unchanged, and is the one stated by Democritus.

Suggestions for the Teacher

(a) Review the process of "hypothesizing." Students should realize that the problem is unchanged, hence the hypothesis will probably remain unchanged.

5. Design for Collection of Data.

   The important aspect here is that the new apparatus, techniques, and procedure should overcome the limitations in the design in Investigation A- 1. The aspect of controlling a variable is introduced (in this case the absence of sugar giving a negative test for comparison).

Suggestions for the Teacher

(a) The design given on page 5 could come from the post-laboratory discussion of Investigation A - 1 (see "further evidence"). Look at the aspects of design again, and illustrate each with appropriate part of the design.

(b) Materials required for the experiment: sugar cube, two 150 ml beakers, water, filter paper, funnel, test tube, sodium hydroxide solution, cobalt nitrate solution.

11. Interpretation of Data

No one has ever seen the basic unit or particle of which sugar is composed. Therefore, observations can be explained only on the basis of inference. Still more refined apparatus and experimental technique may yield better "proof" for the basic unit of matter.

Question 1: The visible sugar disappeared from the filter paper and appeared in an invisible form in the solution. The evidence for this is the positive tests for sugar. We must assume that dissolving action of water breaks down the sugar into particles at least small enough to go through the very small holes in the filter paper.

Question 2: We certainly obtained smaller particles in this experiment than we did in Investigation A - 1. The "holes" in the filter paper were not visible even with an ordinary microscope, yet were big enough to let sugar particles through.

Question 3: Democritus' hypothesis was confirmed only if we assume that the dissolving results in the breakdown of sugar crystals into the smallest particles possible. We have no direct evidence that the basic unit from which visible sugar particles are built was obtained in our experiment.

Question 5: Sugar in the cube is "solid," "visible," and consists of fairly large particles. Sugar in the solution is "liquid", invisible, and may consist of the smallest possible units.

Suggestions for the Teacher

(a) Review the purpose of "hypothesis" and its confirmation or rejection (see answer to question 3).

(b) Note the limitations of the data obtained since an explanation is not self-evident. It has to be inferred or assumed.

12. Operational Definitions

The word used to represent the basic particle or unit in matter is "molecule". The act or event of a substance disappearing in a liquid is labelled "dissolving." In terms of the molecular theory, the explanation for the dissolving process is that individual molecules are freed from the crystal into the liquid.

Suggestions for the Teacher

(a) Discuss the basis for "operational definitions".

(b) Considerable confusion exists in differentiation between an atom and a molecule. The term molecule is used to designate the basic unit of any pure substance (element or compound). In a few substances the molecules are single atoms (e.g., helium). In many the molecules are made up of combinations of atoms of a single element (e.g., oxygen, O_2; sulfur, S_8; phosphorus, P_4). In most substances the molecules are made up of combinations of atoms of two or more elements (e.g., water, H_2O);cane sugar, $C_{12}H_{22}O_{11}$).

14. Development of a Theory

<u>Suggestions for the Teacher</u>

(a) Discuss the idea of a theory, and how it develops. Use the "kinetic-molecular theory" as an example, and show how so far only the idea of molecules is part of it. This theory will be expanded and refined throughout the remainder of the course on the basis of investigating new related problems.

15. Further Evidence

Further evidence for the molecular nature of matter is desirable. To get it for sugar would require still more sophisticated apparatus and techniques. In addition to the property of dissolving which was the basis of this experiment, other properties such as diffusion, ionization, states, etc. could be used. To generalize the concept of molecules to all pure substances would require repeating the same experiments for each one of them. Obviously, this is an impossible task so the scientist uses only a representative selection of substances to test the universality of the idea of molecular structure.

<u>Suggestions for the Teacher</u>

(a) The ideas of students for further evidence should be discussed in class and assessed critically. They can also be incorporated tentatively into the "kinetic-molecular theory" as hypotheses.

(b) The teacher might use appropriate films and film loops to provide further evidence for molecules.

16. New Problems for Investigation

Some examples of new problems are:

- size of molecules (how small?)
- shape of molecules
- numbers of molecules involved
- diffusion (the entire solution became sweet)
- rate of colour change versus concentration of sugar
- structure of filter paper

Suggestions for the Teacher

(a) Discuss new problems in terms of studying new variables)e.g., size of molecules), incidental problems (e.g., structure of filter paper), etc.

(b) Students might be allowed to seek background information and design an experiment for one of the new problems.

(c) Provide guidance in locating suitable reference material for part 2.

(d) Someone might come up with a single method of measuring the size of a molecule (e.g., Experiment I - 8 in P.S.S.C. Physics).

FILM LOOPS USED IN THIS STUDY

1. The Knife - from Suchman's Inquiry Development Program (62)
   University of Alberta Education Library

2. Heating Liquids - Edmonton Public School Board

3. Mixing Liquids - Specially prepared by author

4. The Shrinking Balloon - from Suchman's Inquiry Development Program (62)
   University of Alberta Education Library

5. The B.P. of Water - Edmonton Public School Board

6. Melting Ice Cubes - Prepared by author

7. The ice cubes - from Suchman's Inquiry Development Program (62)
   University of Alberta Education Library

8. The Restaurant - from Suchman's Inquiry Development Program (62)
   University of Alberta Education Library

9. Balancing the Ball - Edmonton Separate School Board

APPENDIX B

MATTER AND ENERGY

1. When air in a closed vessel is cooled,

   A. the volume of the air diminishes
   B. the sizes of the molecules decrease
   C. the molecules move less rapidly
   D. the spaces between the air molecules decrease

2. Some boys took bottles of pop on a hike up a high mountain. At the top they noticed the level of the pop in the bottles was lower. The reason for this was that the

   A. atmospheric pressure is higher the higher you go
   B. atmospheric pressure decreases with altitude
   C. temperature was lower at higher altitude
   D. pressure inside the bottles increased

3. Heat may be defined as

   A. the total kinetic energy possessed by all the molecules in a body
   B. the total temperature of a body
   C. the total temperature of a body at a certain pressure
   D. the number of molecules in a body

4. Temperature is defined as a measure of the

   A. number and speed of molecules of a substance
   B. number of molecules in a substance
   C. average kinetic energy of molecules in a substance

5. Gases are poor conductors of heat because their molecules are

   A. smaller than those of solids or liquids
   B. moving about too rapidly
   C. not freely moving but can only vibrate in a lattice
   D. too far apart to transfer energy by collision

6. When a gas in a closed container is heated, the pressure increases; and when it is cooled, the pressure decreases. This is because

   A. gas molecules expand, thus taking up more volume
   B. gas molecules move about faster when heated, thus colliding more often with the walls of the container
   C. the inner space of the container increases upon heating
   D. the closed container is a metal and is therefore a better conductor of heat than a gas

7. Gas thermometers are more sensitive to small temperature changes than either liquid or bi-metallic thermometers because

   A. the molecules in gases are farther apart than in liquids or solids
   B. gases expand and contract more than equal volumes of liquids and solids for the same temperature change
   C. gases are the best convectors of heat
   D. gases absorb heat when they expand but give off heat when they contract

8. Expansion of a solid is the result of

   A. a lattice structure of the atoms
   B. atoms increasing their kinetic energy
   C. the space between atoms and molecules
   D. molecules occupying more space than atoms

9. If a large quantity of cold water is quickly poured into the radiator of a very hot car engine, the engine block may crack because

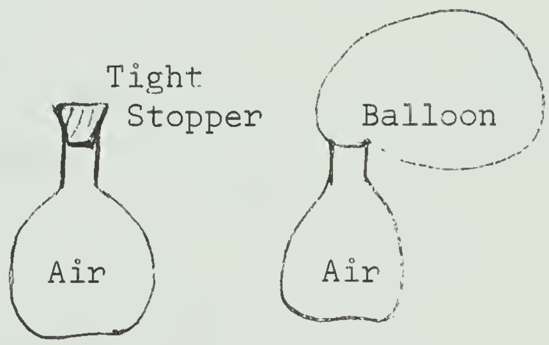
   A. the water produces steam under high pressure when it comes in contact with the hot engine
   B. automobile engines are designed to operate at high temperatures
   C. the cold water cools the parts of the engine which it contacts and other parts remain hot
   D. the cold water expands when it gets inside the hot engine block

10. If a can of beans is heated strongly before being opened, it will burst because

   A. a chemical reaction inside produces carbon dioxide
   B. the heat weakens the soldered joints
   C. there is an increase in the molecular movement of the contents
   D. the molecules inside the can will expand

11. Two identical flasks are filled with air, one with a balloon over the top and the other sealed tight. If they are both placed on the same hot plate, in which one is the temperature rise slower?

   A. in the sealed flask because the gas cannot expand
   B. in the balloon topped flask because the molecules can expand
   C. in the sealed flask because the molecules cannot expand
   D. in the balloon topped flask because the air can expand



12. The temperature at which a liquid boils may be raised by

    A. insulating the container
    B. covering the container with an airtight cover
    C. removing the vapour from the surface above the liquid
    D. changing the container from a metal to glass

13. Which of the following will NOT increase the evaporation rate of a liquid?

    A. increase the temperature of the liquid
    B. cause the air above the liquid to circulate
    C. increase the atmospheric pressure
    D. increase the surface area of the liquid

14. Evaporation from a liquid results in

    A. a decrease in the average speed of the molecules in the liquid
    B. an increase in the temperature of the liquid
    C. an increase in the rate of collisions between molecules of the liquid
    D. an increase in the average speed of the molecules of the liquid

15. In the process of evaporation cooling takes place because

    A. molecules of the vapour have less kinetic energy than those of the liquid
    B. during evaporation, the molecules with the greater kinetic energy escape from the liquid
    C. water vapour conducts heat rapidly
    D. evaporation makes the air damp

16. If the watchglass containing oil and water was left to sit in a classroom for three hours, what would you predict would probably happen?

    A. Both liquids would evaporate the oil more than the water.
    B. Both liquids would evaporate the water more than the oil.
    C. Both liquids would evaporate the same amount.
    D. The oil would evaporate slightly and the water not at all.

17. In a submerged submarine, water in the galley (kitchen) boils at a temperature of

    A. over 100 degrees centigrade because pressure is reduced
    B. over 100 degress centigrade because air pressure is increased
    C. over 100 degrees centigrade because the water pressure increases with depth
    D. near 100 degrees centigrade because the air pressure in the kitchen probably has not changed much from what it was above sea level.

18. On a very hot day an electric fan produces a cooling effect by

    A. bringing in cold air and removing warm air
    B. directing the hot air outside
    C. increasing the rate of evaporation
    D. increasing the circulation

19. When a substance is heated, its molecules

    A. expand
    B. contract
    C. start vibrating
    D. increase their motion

20. Molecules are closest together in

    A. liquids
    B. solids
    C. gases
    D. fluids

21. Diffusion occurs most rapidly between a

    A. solid and a liquid
    B. liquid and a liquid
    C. gases
    D. fluids

22. When a solid dissolves in a liquid, its molecules

    A. become so small they disappear
    B. move freely among the molecules of the liquid
    C. dissolve
    D. none of the above

23. If a vinegar bottle is left open in a large room, a vinegar smell can be detected after some time anywhere in the room because

    A. vinegar has a very strong smell
    B. vinegar does not evaporate readily
    C. vinegar molecules diffuse into the air and occupy the lower layers of the room
    D. vinegar molecules diffuse into the air and fill the whole space available to them

24. Two liquids with characteristic odors are placed in opposite corners of a room. If you stood half way between the open bottles containing the liquids,

    A. you could not smell any of the two liquids
    B. after some time you could smell the liquid that is made up of the heavier molecules
    C. after some time your sense of smell could probably detect both liquids
    D. if two liquids are used, neither one of them will diffuse to any great extent

25. When a liquid freezes

    A. its molecules move more slowly
    B. its molecules decrease in size
    C. it occupies less space than the liquid
    D. it must take in heat energy

26. As a pure substance melts, its temperature

    A. falls rapidly
    B. rises rapidly
    C. rises slowly
    D. remains the same

27. Two changes of state that a substance can undergo only by giving up heat are:

    A. condensation and freezing
    B. melting and vapourization
    C. evaporation and boiling
    D. melting and freezing

28. To prevent sugar from scorching it is boiled in a vacuum pan because

    A. boiling takes place at a lower temperature
    B. increased pressure causes faster evaporation
    C. burning requires the presence of oxygen
    D. boiling takes place at a higher temperature

29. Edmonton has an elevation of about 2,000 feet, Calgary about 3,000 feet and Montreal about 400 feet. The boiling point of water will be

   A. lowest at Montreal, highest at Calgary
   B. roughly the same at all three cities
   C. highest at Montreal, lowest at Calgary
   D. lower at Montreal, Calgary and Edmonton roughly the same

30. A gas differs from a solid or a liquid because a gas

   A. expands to fill its container
   B. has no definite shape
   C. is weightless
   D. has size of its own

31. A property possessed by all matter is

   A. a definite shape
   B. a definite volume
   C. weight
   D. no definite size

32. The molecular theory of matter

   A. is based on assumptions only with no experiments being made
   B. contains assumptions that are accepted because they agree with observations obtained from experiments
   C. cannot be considered true because molecules cannot be measured or directly seen
   D. probably will never be altered because it explains changes of state so well

33. Heat energy is released when

   A. moth balls "disappear" into the air
   B. liquid sulphur hardens to solid sulphur
   C. alcohol evaporates into the room
   D. alcohol reaches its boiling point

34. The most regular arrangement of molecules can probably be found

   A. in a melting solid
   B. in a liquid
   C. in a solid
   D. in a gas

35. Convection occurs only in

    A. non-metals
    B. liquids
    C. gases
    D. fluids

36. The fact that some substances dissolve in others without an increase in the overall volume suggests that

    A. a weight loss probably occurs
    B. there are spaces between molecules
    C. molecules are tightly packed
    D. our instruments for measuring volume are not sensitive enough

37. When a tire is pumped up, molecules are

    A. heated
    B. contracted
    C. added
    D. expanded

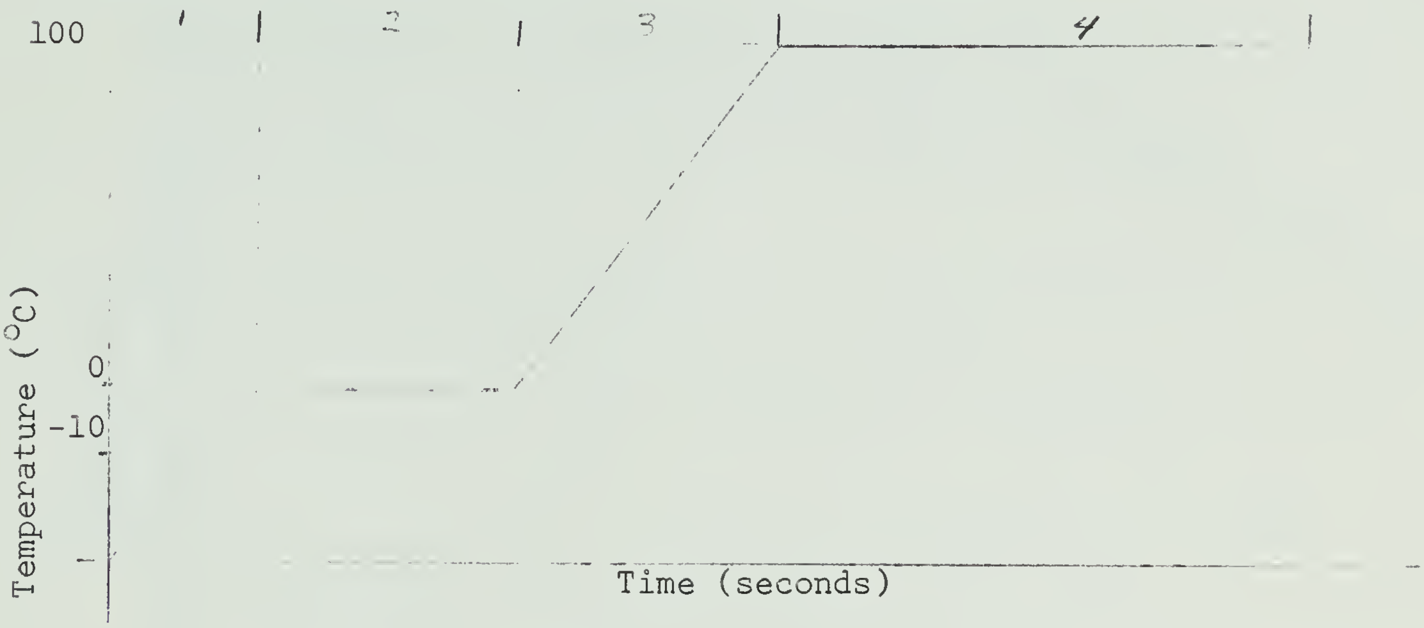
38. A pupil glued some marbles together to make a model of how molecules are arranged in a solid. This model suffers from a number of limitations, the most important of which is probably the fact that

    A. all molecules are not usually the same size
    B. the space between the molecules compared to the size of the molecules is usually larger
    C. molecules in a solid are not arranged as regularly as in the marble model
    D. molecules are much more solid than marbles are

39. A pupil wants to get a can as full of molecules as possible. In order to do this he would obtain a sample of air from

    A. a mountain top
    B. a beach
    C. from Edmonton
    D. the stratosphere

Questions 40 - 42 are based on the following graph:



This graph summarizes what happens to a quantity of ice as it is heated to its boiling point. Now answer questions 40 - 42.

40. The portion of the curve where the volume change is the greatest is

A. 1
B. 2
C. 3
D. 4

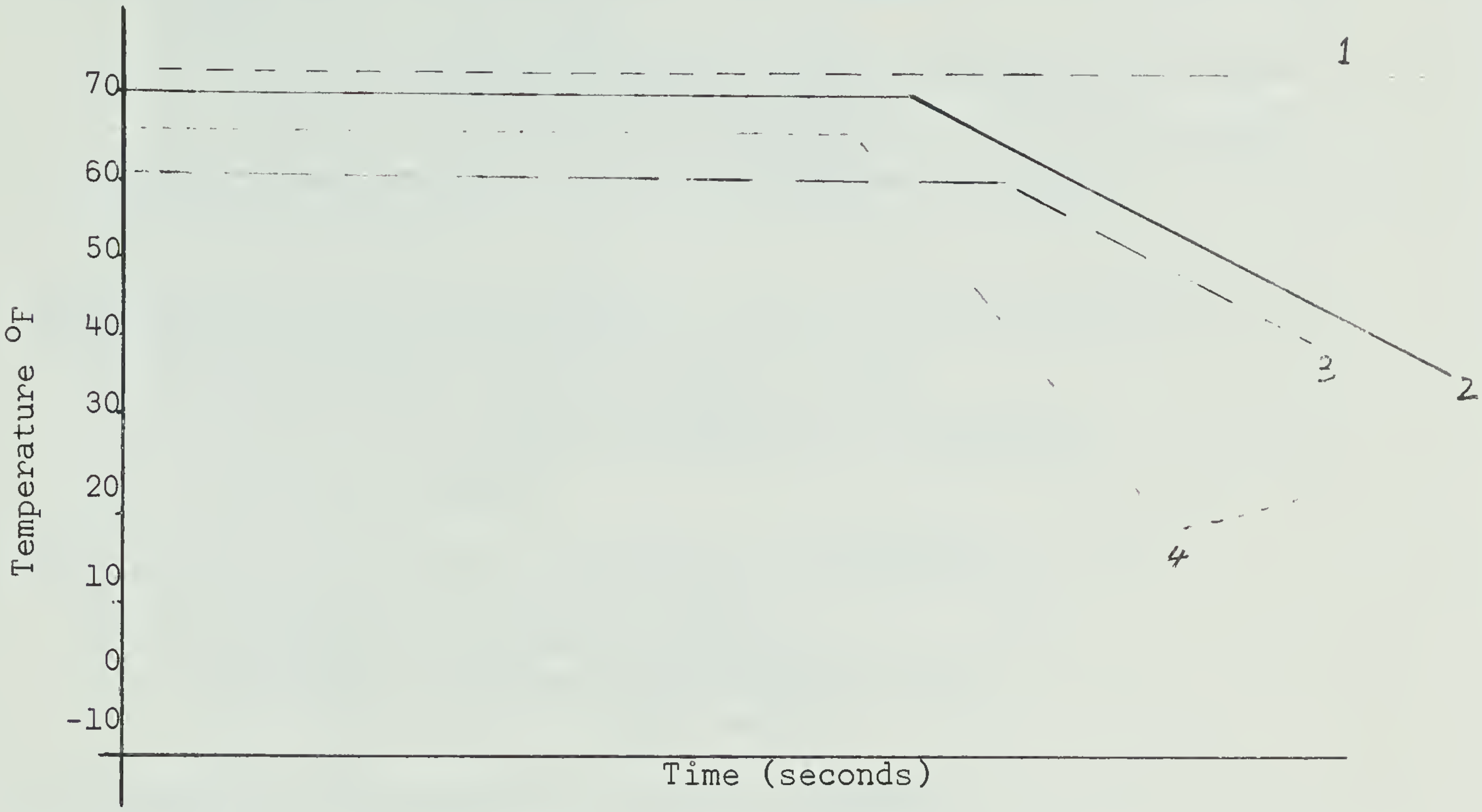
41. The part of the curve where molecules have the least energy is

A. 1
B. 2
C. 3
D. 4

42. The portion of the curve where no temperature change occurs is

A. 1 and 2
B. 2 and 3
C. 2 and 4
D. 1 and 3

Questions 43 - 45 are based on the following graph:



The above is taken from a student's notes based on an evaporation experiment. The graph shows what happens to the temperature of four different liquids as they evaporate.

43. The liquid that evaporated fastest was

A. 1
B. 2
C. 3
D. 4

44. The liquid that has the highest boiling point is probably

A. 1
B. 2
C. 3
D. 4

45. The liquid that evaporated completely during the experiment was

A. 1
B. 2
C. 3
D. 4

PART I - THE KNIFE

1. What was the main hypothesis upon which this experiment is based?
   A. If the knife is heated and then dipped in a liquid, then the blade does not get black.
   B. If a knife is heated and then cooled, the knife can be made to bend and then straighten out again.
   C. If a knife is heated and then dipped in a liquid, the knife can be made to bend and then straighten out again.
   D. All of the above.

2. What do you predict would happen if the knife was placed in an ice bath instead of being heated?
   A. The molecules in the knife could freeze.
   B. It would bend the same way than if it was heated.
   C. It would bend opposite to the knife in the film.
   D. It would get shorter.

3. Why was the knife caused to bend both up <u>and</u> down?
   A. To show that it was not a trick.
   B. To eliminate the effect of gravity.
   C. To show that heat bends metal only one way.
   D. To show that heat can expand metals.

4. Why was the knife wiped off after each dipping?
   A. To prevent rust.
   B. To prevent the dripping from putting out the fire.
   C. To prevent steam from forming.
   D. To remove the effect of the liquid during heating.

5. You probably observed a smooth rather than an irregular bending of the knife. What two factors are likely responsible for this?
   A. The heat and carefulness of heating.
   B. The wide flame and even heating.
   C. The temperature of the flame and the temperature of the liquid.
   D. The temperature of the flame and heating in one spot.

6. What do you think caused the bending of the knife?

   A. The melting
   B. The heat of the flame softened the metal and the liquid hardened.
   C. The moisture of the liquid and the heat of the flame.
   D. None of the above.

7. What purpose, besides providing a suitable place to hold the knife, did the handle serve?
   A. To give the expanding metal a place to begin expanding.
   B. To prevent the man from cutting himself.
   C. To keep the knife from bending too far.
   D. To show when the knife is straight.

8. What useful purpose could a device such as this serve?
   A. It shows that things are easier to cut when the cutting instrument is cold.
   B. It could be used to show that the knife is made of stainless steel.
   C. To indicate a change in temperature.
   D. It can be used to show that heat can bend anything.

9. How would the fact that the knife was held vertically above the flame instead of horizontally affect the results?
   A. the knife would not bend.
   B. There would be no difference in results.
   C. The bending would be sideways rather than up and down.
   D. The knife would just get longer.

HEATING LIQUIDS

10. What is the hypothesis upon which this experiment is based?
   A. If a liquid is heated, then its temperature will rise.
   B. If a liquid is heated, then its volume increases.
   C. If a liquid is heated, it begins to weigh less.
   D. If a liquid is heated, then it begins to weigh more.

11. What is the significance of the level not returning to "A"?
   A. Once a liquid expands, it cannot contract to original size.
   B. The liquid must have become too cold.
   C. The pressure must have increased.
   D. The temperature was still higher than it was at "A".

12. What does the fact that the scales did not balance indicate?
   A. Cold water weighs more than hot water.
   B. If equal volumes are used, cold water will outweigh warm water.
   C. The volumes of H_2O must be different.
   D. A factor not shown affected the balance.

13. What additional variable would you introduce if you stoppered the thin tube in which the liquid rises?
   A. pressure
   B. steam
   C. condensation
   D. the liquid would not rise.

14. If you suspected some trickery or error in weighting, a very simple technique used with the double-pan balance is to
   A. weigh two objects known to have equal weight.
   B. interchange the two beakers
   C. re-balance the empty scale with empty beakers
   D. any of the above

15. What happens to the weight of the flask as the liquid in the flask is heated and increases in volume?
   A. It gets heavier
   B. It gets lighter
   C. It gets heavier because the volume increases.
   D. It remains the same.

16. Why was the large flask filled completely before the stopper was inserted?
   A. To remove all the air so that a change in volume of the liquid shows easily.
   B. A certain weight of water was required.
   C. So the water cannot turn into steam.
   D. The water expands faster under these conditions.

17. How does the radiator in the modern automobile recognize the principle under investigation?
   A. It has an airtight cap to make the water boil faster.
   B. It has an airtight cap to keep the water from expanding.
   C. It has an overflow tube to allow for expansion.
   D. It is always filled up completely to avoid pressure.

18. If you were graphing the results in the experiment with the colored liquid, which two variables would you plot against each other?
   A. Temperature against pressure of the air.
   B. Temperature against volume of the liquid.
   C. Temperature against weight of liquid.
   D. Temperature against pressure in liquid.

MIXING LIQUIDS

19. What question does the experiment answer?
   A. Do water and alcohol react at room temperature?
   B. If water and alcohol are mixed at room temperature can the loss in weight be accurately determined?
   C. Is the total combined volume of two liquids the same as the sum of the separate volumes?
   D. If water and alcohol are mixed, is heat evolved?

20. What hypothesis is this experiment based on?
   A. If water and alcohol are mixed in the right porportion, then a reduction in the total combined volume occurs.
   B. If equal amounts of water and alcohol are mixed, then a reduction in total combined volume occurs.
   C. If water and alcohol are mixed on a balance, then a loss of weight can be observed.
   D. If water and alocohol are mixed on a balance, then a gain in volume occurs.

21. Why were two 250 m. flasks of H_2O poured together first?
   A. To show that when liquids of the same kind are mixed, then no volume loss occurs.
   B. To show that the jars were actually 250 ml. each.
   C. To show that the measured volumes were accurate.
   D. To show that when liquids of the same kind are mixed, no weight change occurs.

22. Why was the last part of the experiment performed on a two-pan balance.
   A. To show that the two small flasks weighed the same as the one empty large one.
   B. To show that a loss of weight occurred.
   C. To show that a gain in weight occurred.
   D. To show that no change in weight was found.

23. What did you observe when the water and alcohol were poured together?
   A. There was no volume change.
   B. There was a gain in volume.
   C. There was a loss in volume.
   D. There was a loss in weight.

24. A pupil performed the same experiment but used only 5 ml. of each liquid. His results were not as conclusive as those in the film. Why not?
   A. He should have repeated the experiment many times.
   B. He did not use a weigh scale.
   C. He did not use large enough volumes to indicate a clear difference in weight.
   D. He did not use large enough volumes to indicate a clear difference in volume.

25. What does the total combined volume of two different liquids like water and alcohol tell us about the arrangement of molecules in these liquids?
   A. The molecules are all the same size.
   B. There are spaces between the molecules.
   C. The molecules are probably of different shape.
   D. The molecules are stacked tightly together.

26. If one wanted to build a model of what occurred in the experiment, one would pour
   A. two sets of the same-size marbles together.
   B. sand and gravel together.
   C. yellow sand on white sand.
   D. salt and sugar together.

27. An important application of the principle under investigation is
   A. that liquids expand when heated
   B. that water contracts when freezing
   C. that gases can be compressed.
   D. all of the above.

PART II - THE SHRINKING BALLOON

1. What is the problem under investigation?
   A. What happens to a balloon when it is heated?
   B. How does temperature affect the volume of a gas?
   C. Can the shape of a ballon be changed?
   D. Can a balloon shrink when heated and expand when cooled?

2. What happened to the molecules in the balloon when it was lowered into the jar?
   A. They escaped from the balloon.
   B. They contracted.
   C. They slowed down.
   D. They stopped moving.

3. How is the action of the balloon similar to that of an automobile tire during a hot summer day?
   A. Both get flat upon heating.
   B. The tire does not get as flat as the balloon.
   C. Both expand on heating.
   D. Both get hot and become soft.

4. What is the hypothesis upon which this experiment was based?
   A. If the temperature of the air in a balloon is changed, then the volume will change.
   B. If a balloon is heated, then it loses weight.
   C. Cooling a balloon can cause it to expand back to the original size.
   D. all of the above

5. What do you predict will happen if the balloon were held over a steam bath at the end of the film?
   A. It would contract.
   B. It would expand and burst.
   C. It would expand.
   D. It would expand and rise.

6. With which factor was the <u>main</u> observation concerned?
   A. The weight of the balloon.
   B. The volume of the air in the balloon.
   C. The temperature and pressure of the surroundings.
   D. The weight of the air in the balloon.

7. What was the main reason for using the balloon in this experiment?
   A. It changes shape rapidly.
   B. It does not burst when heated.
   C. It shows a change in volume readily.
   D. It is easily handled.

8. If you were to make actual measurements,which of the following measurements would be the most useful?
   A. How long it takes to shrink and expand the balloon.
   B. The volume before and after putting the balloon into the jar.
   C. The weight of the balloon before blowing it up.
   D. The air pressure during the experiment.

9. What was the main purpose of the little "board" on which the balloon was resting?
   A. To expand the balloon.
   B. To support and insulate the balloon.
   C. It acted as a source of cold.
   D. It allowed the balloon to cool off.

THE BOILING POINT OF WATER

10. Which one of the following questions is answered by the experiment?
   A. Can the same water be made to boil by various means?
   B. Can ice cool water enough to make it boil?
   C. Can the boiling point of water be made to vary?
   D. Does water expand when heated?

11. What do you predict would happen to the water if the ice was put on the flask after the water had cooled off for the better part of a day?
   A. The water would get colder.
   B. The water would begin to boil again.
   C. The water would freeze.
   D. Nothing would happen.

12. Why was the ice pack not applied to that part of the flask containing the liquid?
   A. The flask might burst.
   B. The ice would melt too fast.
   C. The liquid would boil vigorously.
   D. The liquid would cool and not boil.

13. After closing the clamp
   A. The boiling point of water cannot be varied to any large extent.
   B. The ice should not be applied too fast.
   C. The pressure in the flask has to remain constant.
   D. The liquid in the flask should not be boiled for too long.

14. How long could boiling be continued after the ice is placed on the flask?
   A. indefinitely.
   B. as long as there is ice.
   C. for a short time only until the temperature of the water gets too low.
   D. as long as molecules are moving.

15. Besides not allowing one to see what is happening in the flask, why would a copper container be a poor choice for this experiment?
   A. It could melt at these high temperatures.
   B. It conducts heat too well.
   C. It cannot stand large temperature changes.
   D. It cannot stand large pressure changes.

16. The latter half of the film shows that there is a definite relationship between two variables. The two are?
   A. volume and temperature.
   B. boiling point and temperature.
   C. boiling point and pressure.
   D. volume and pressure.

17. Water on the surface of the earth rarely has to be heated higher than 100 degrees Centigrade to boil. What is the reason for this?
   A. Normally the pressure is not high enough to cause it to boil at a higher temperature.
   B. Most heat sources cannot heat beyond 100 degrees C.
   C. Water always boils at 100 degrees C.
   D. It can boil at a lower temperature but not a higher.

18. Boilding an egg on a mountain top is similar to one of the situations in the film. The situation is
   A. boiling water in an open flask.
   B. boiling water after pumping air into the flask.
   C. boiling when an ice pack was applied to the flask.
   D. boiling under extreme pressure.

MELTING ICE CUBES

19. What problem is being investigated here?
   A. Does ice melt before lead?
   B. Can water boil before ice in it melts?
   C. How well does water conduct heat?
   D. Can a lead shield keep heat from ice cubes under it?

20. Why was it wise to have two test tubes?
   A. It shows that the rate of melting depends where the tube is heated.
   B. In case one test tube did not show the desired result, the other one probably would.
   C. If both showed the same result, then the conclusion could not be questioned.
   D. None of the above.

21. What is the purpose of the lead plate?
   A. To prevent heat from getting to the ice cubes.
   B. To hold the cubes on the bottom of the tubes.
   C. To absorb the excess heat in the water.
   D. Lead was used because it does not react with the water.

22. Another pupil performed a similar experiment but he let an ice cube float and showed that it melted if the test tube was heated on the bottom. Why are the two experiments not the same?
   A. Both convection and conduction are being studied here.
   B. Water does not conduct heat vertically.
   C. Ice would melt at the top eventually but not at the bottom.
   D. There was no use for lead here.

23. Why is the glass preferable to metal containers?
  A. Glass can withstand a greater temperature change.
  B. Glass can be heated in one spot more safely.
  C. Glass does not conduct heat as well.
  D. Glass heats more evenly than metals.

24. The tubes could probably have been best suspended vertically. Why were they at an angle?
  A. Water conducts heat better this way than vertically.
  B. To make it easier to heat one tube near the top of the water level.
  C. It is not too safe to heat a test tube suspended vertically.
  D. To allow the steam to escape more easily.

25. Rate of melting is definitely a variable in this investigation. Which additional factor could cause a change in the rate of melting?
  A. How thick the lead plate is.
  B. The pressure of the air.
  C. The temperature of the air.
  D. The size of the flame.

26. What can you say about the ability of water to conduct heat?
  A. It probably conducts heat as well as any other liquid.
  B. It is a poor conductor of heat.
  C. It is a good conductor of heat.
  D. Cold water does not conduct heat as well as warm water.

27. Nature is full of examples of the way in which water conducts heat. Which one of the following is <u>not</u> an exmpale of this property?
  A. One can sit in one end of the bathtub with a red-hot iron in the other end.
  B. Ocean water can be both warm and cold in one area at the same time.
  C. Hot water heating is more efficient than hot air heating.
  D. Water is often used to cool an automobile engine.

PART III- THE ICE CUBES

1. What problem is being investigated here?
   A. Do ice cubes melt faster in hot or cold water?
   B. Do the ice cubes float in all liquids?
   C. What is the effect of cube size on its ability to float?
   D. Can an ice cube be made to sink after floating in a liquid?

2. Before any definite conclusions can be drawn, what additional useful tests might you perform?
   A. Measure the size of the cubes.
   B. Test the liquids to see if they are the same.
   C. Take the room temperature.
   D. Take the temperature of the cubes.

3. Which factor was controlled rigidly in this experiment?
   A. The temperature of the liquids.
   B. The size of the cubes.
   C. The kind of liquids.
   D. The amount of liquids used.

4. If you wanted to make actual measurements, what two measurements might be the most useful?
   A. Temperature and weight of each liquid.
   B. Temperature and size of cubes.
   C. Size of cubes and the air temperature.
   D. Size of cube and time it takes to melt.

5. What would be the clearest way to record the observations in this experiment?
   A. Make a table.
   B. Draw a diagram.
   C. Draw a bar graph.
   D. Draw a line graph.

6. What is the most general conclusion one can draw from the experiment evidence?
   A. The ice cools hot water.
   B. Cubes at different temperatures do not float the same.
   C. The cubes react with the liquids.
   D. The liquids must be different in some way.

7. Time could become an important factor in this experiment. What observation has led to this statement?
   A. One cube started rising and the other sinking.
   B. One cube started to rise.
   C. Both cubes melted very quickly.
   D. The stirrer made one cube rise while it did not affect the other.

8. In a similar experiment a pupil was given two equally full glasses of distilled water and two cubes. Inserting the cubes in the glasses gave the same observations as the filmed experiment. Which factor probably differed in this case?
   A. The amount of liquids.
   B. The type of liquid.
   C. The material the cubes were made of.
   D. The temperature of the cubes.

9. What is the similarity between this experiment and swimming in the ocean right after swimming in a fresh water pool?
   A. One sinks if he swims long enough--even in salt water.
   B. The body acts like a cube and only the liquids differ.
   C. It is easier to swim in the ocean because it is usually colder.
   D. None of the above.

THE RESTAURANT

10. Which of the following questions is answered by the experiment?
   A. Is the hand quicker than the eye?
   B. Does the success of pulling a tablecloth from underneath dishes depend on the speed of action?
   C. Does the success of pulling a tablecloth from underneath dishes depend on who is doing it?
   D. Can a tablecloth be pulled from underneath a table setting without disturbing a person while he is eating?

11. The best hypothesis upon which this experiment is based is
   A. Experience is necessary to perform tricks properly.
   B. If a tablecloth is yanked from underneath dishes with a snappy action, then the dishes will not be upset.
   C. If the proper dishes are used, then one can perform this trick quite easily.
   D. If two people try this trick, only one is likely to succeed.

12. What do you predict would happen if waxpaper instead of linen were used for a "cloth"?
   A. The dishes would slip too easily disrupting the setting.
   B. The waxpaper would probably tear.
   C. Waxpaper would be harder to pull out safely.
   D. Waxpaper is easier to pull out safely.

13. Why did the waiter have more success than the customer?
   A. The waiter had a steadier hand and more practice.
   B. The waiter pulled the tablecloth out much faster and "snappier".
   C. The waiter had his table set up for the trick.
   D. The waiter did not let the customer see everything there was to the trick.

14. Which was the factor that changed most between the trials?
   A. The number of objects on the table.
   B. The weight of the objects on the table.
   C. The type of the cloth used.
   D. The speed of pulling the tablecloth.

15. There is probably only one measurement that might be useful in this experiment. In order to carry out this measurement you would probably need an instrument that measures:
   A. area of table.
   B. weight of dishes.
   C. speed.
   D. time.

16. If you were doing this experiment at home, what changes would you make in the design?
   A. Use some unbreakable dishes and arrange them evenly on the table.
   B. Use a smaller table.
   C. Iron the cloth before doing the experiment.
   D. Use some unbreakable dishes and contents that do not spill.

17. If you thought there was more than the <u>one</u> factor acting in this experiment, how would you change your experimental design to show this?
   A. You would set up more tables.
   B. You would get more people to participate.
   C. You would use a variety of speeds.
   D. No change in design.

18. In movies, the criminals are often disarmed by "getting the rug pulled from under their feet." What is the important difference between this situation and the filmed experiment?
   A. A human being takes the part played by the dishes.
   B. In one case it is intentional to upset objects.
   C. It is much easier to pull a rug from under a person than to pull a tablecloth from underneath dishes.
   D. There is no basic difference.

<u>BALANCING THE BALL</u>

19. Which of the following questions is answered by the experiment?
   A. How does pressure affect weight?
   B. Does air have weight?
   C. Does temperature affect weight?
   D. Can a long stick be used as a balance?

20. What do you predict would happen if after the balance was equalized, the ball was heated over a steam bath?
   A. The ball would get larger and gain in weight.
   B. The ball would get larger but the balance would remain the same.
   C. The ball would remain the same size and the balance remain the same.
   D. The ball would remain the same size but get heavier.

21. Why was the slider on the balance adjusted at the start of the experiment?
   A. To balance the stick with a partially filled ball at one end.
   B. To make sure it was not stuck.
   C. To show that it could be used to balance the scale if necessary.
   D. None of the above.

22. What did you observe when the ball was pumped up and replaced on the balance?
   A. The ball rose in the air.
   B. The ball became larger but weighed the same.
   C. The ball weighed more.
   D. The ball became larger but weighed less.

23. You probably noticed that the experimenter used fine grains of material to achieve a balance. Why?
   A. No standard weights were available.
   B. Each grain weighs the same.
   C. The number of grains give an indication of the weight in grams.
   D. The grains can be used to achieve a very accurate and quick balance.

24. What purpose does the scale which was introduced at the end serve?
   A. It shows actual weight.
   B. It shows the amount of deflection and weight in grams.
   C. It shows the point of balance and amount of deflection.
   D. None of the above.

25. If no grains nor other weights were available to balance the scale after a deflection, how could a balance be achieved?
   A. By adjusting the stick so that it balances.
   B. By moving the slider.
   C. Either of the above.
   D. Either one of the above two but not both at the same time.

26. What general conclusion can you reach from the experimental evidence?
   A. Air under pressure has weight.
   B. Air has weight.
   C. An empty ball outweighs a ball filled with air.
   D. The weight of air can be determined by using a simple balance.

27. Two empty balls are of equal weight. They are of equal size when filled with air, yet do not reach the same balancing levels. What can be concluded?
   A. The volume of air in one was larger than in the other.
   B. Something went wrong with the balance.
   C. One contained more air than the other.
   D. One was more buoyant than the other.

TRIAL RUN

(CHECKING ON THE EFFECTIVE ERROR VARIANCE TO BE USED FOR THE DETERMINATION OF THE SIGNIFICANCE OF DIFFERENCE AMONG ADJUSTED TREATMENT MEANS, SEE WINER (73) p. 620)

TREATMENT	A		B		C		D	
TEST Part I of the Process Measure	COOP(X)	PRETEST(Z)	COOP(X)	PRETEST(Z)	COOP(X)	PRETEST(Z)	COOP(X)	PRETEST(Z)
	46	12	36	16	29	13	26	7
	44	14	40	9	21	12	41	12
	35	20	37	13	40	10	27	7
	46	15	27	11	49	11	43	11
	32	12	41	8	43	17	19	7
	36	14	29	11	33	11	24	11
	36	12	38	16	47	12	18	7
	52	16	37	10	31	15	24	11
	38	10	37	17	39	13	15	12
	37	12	38	15	28	9	27	9
	46	10	44	13	41	22	45	16
	21	8	39	17	41	14	45	15
	23	9	42	9	17	10	35	14
	39	11	37	12	22	14	38	17
	14	6	43	11	34	11	40	16
	13	6	28	14	20	16	54	18
	27	10	39	18	44	16	37	10
	29	13	35	13	22	11	50	14
	44	12	45	19	23	8	26	15
	35	15	41	12	33	11	23	8
ΣX; ΣZ	693	237	753	264	657	256	657	237
ΣX^2; ΣZ^2	26,229	3025	28,817	3680	23,425	3478	24,035	3059
ΣXZ	8660		9964		8591		8333	

$$\frac{T_{xx}E_{zz} - 2T_{xz}E_{xz} + T_{zz}E_{xx}}{(k-1)(E_{xx}E_{zz} - E_{xz}^{2})} = \frac{(308)(864) - 2(54)(1201) + (28)(6978)}{(3)\ \left[(6978)(864) - 1,442,401\right]}$$

$$= .024$$